THE COST OF HAPTICS: MEASURING ENCUMBRANCE IN GLOVES FOR MIXED REALITY

A Dissertation
Presented to
The Academic Faculty

by

James Hallam

In Partial Fulfillment
of the Requirements for the Degree
PhD of Architecture in the
College of Design

Georgia Institute of Technology
December 2020

THE COST OF HAPTICS: MEASURING ENCUMBRANCE IN GLOVES FOR MIXED REALITY

Approved by:

Prof. Jim Budd, Advisor
School of Industrial Design
Georgia Institute of Technology

Dr. Wei Wang
School of Industrial Design
Georgia Institute of Technology

Dr. Young-Mi Choi
School of Industrial Design
Georgia Institute of Technology

Dr. Raymond King
Facebook Reality Labs

Dr. Deborah Boehm-Davis
Facebook Reality Labs

Date Approved: November 13, 2020

To the members of the Interactive Product Design Lab, past and present.

ACKNOWLEDGEMENTS

Thanks to my adviser, Prof. Jim Budd, for a decade of support and mentorship – through three degrees and across two schools, my life was literally changed by his decision to help bring interactive technology prototyping to designers. Thanks as well to Dr. Raymond King, who saw the value of this work, and fought to find a home for it (and me), over the last two years. Special thanks as well to the other members of my committee: Dr. Wei Wang, Dr. Young-Mi Choi, and Dr. Deborah Boehm-Davis – for their support and guidance in bringing this work to a meaningful conclusion.

I want to give my gratitude and thanks to Chip Connor and Taylor Bunge – two amazing Research Assistants who helped carry this project through hundreds of hours of facilitation, recruitment, and documentation. Running these studies during a global pandemic was an enormous undertaking, and this project would not have been possible without their efforts. Additionally, huge thanks are due to Mary Ellen Berglund, who led the fabrication team making the custom gloves for this project, and Brian Cox, who contributed so much expertise to the build. Also, to my colleagues at the Interactive Product Design Lab – thank you for your years of friendship and support.

Finally, my thanks to my family – Dave, Elly, Jon, and Setareh – for their care and love from a distance, even as the two-year plan stretched out to eight. And to my wife Emily – I love you and owe you my unending thanks. We met in the IPDL over long nights of work, and there is no one who could have understood and supported me as you did.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Smart gloves are a category of peripherals that promise to transform how people use their hands in computing systems. These gloves are intended to support the use of hand interactions that closely mimic the way hands are used in the real world – allowing for natural object interactions and, through the use of haptic systems extend the sense of touch into the virtual world. The recent availability of affordable consumer-grade Virtual Reality (VR) systems has accelerated the pursuit of smart glove peripherals that incorporate numerous sensors and actuators that enable these interactions.

While these gloves present a wide array of new enriching interactions to users, they also bear costs, in the form of various degrees of encumbrance. Encumbrance is felt by the user as a temporary disability that limits their hand function and access to the world around them. This limits the user's experience in VR, where the focus is on virtual object interaction, and this will be more challenging with the adoption of Mixed Reality (MR), which requires interaction with both virtual and physical objects. Smart gloves for VR are highly complex, as they require the careful integration of numerous computing/sensing systems, soft-goods product development, interaction design, soft-robotics, and haptics. Making gloves for MR will require even more careful design and testing to ensure that the user has unencumbered access to any object or interface they encounter in both virtual and physical space.

This research examines the costs and trade-offs that users might encounter while using haptic gloves in Mixed Reality. These encumbrances were measured following a mixed-methods, task performance approach, with the goal of recommending design options for

future design development. The methods used were drawn from ergonomic and occupational therapy approaches to measuring hand function, in order to identify clear measures of encumbrance imposed by specific design details of various gloves.

This research developed a comprehensive testing protocol which identified 15 key "Costs of Haptics" which encumber the wearer of VR gloves, in exchange for other benefits. These costs are discussed in detail along with the underlying evidence that supports them and will provide a valuable set of design recommendations that should allow design teams to optimize design decisions in the development of the next generation of MR gloves.

CHAPTER 1. SMART GLOVES FOR MIXED REALITY

Interactive gloves have been in development for decades, from the Nintendo Power Glove to the various recent startups and Kickstarter campaigns pursuing improved glove-based hand tracking and haptic feedback. It is natural to look at introducing gloves into the Virtual Reality (VR)/Mixed Reality (MR) technology space as part of the push for deeper realism and immersion, coupled with natural interfaces. Gloves seem like a reasonable way of delivering this new technology to the user – they are easy to understand, conform well to our hands, and we have a long history of building them. Yet there is also a lot we do not understand about how interactive gloves will work for people in their everyday lives.

Gloves are common garments, but not typically worn outside of specific contexts – they are rarely seen worn around the home or office, for example. Gloves have been adopted primarily when they fit in one of these broad categories: barrier protection, thermal protection, fashion, and functional augmentation. There may be a few instances that fall outside these categories, but generally we always have a specific reason to be wearing them. Gloves have come in and out of fashion over the years – they were considered essential for daily wear as recently as 1900 – but they are now not typically worn unless we have a specific functional reason to do so [11].

Understanding this context of daily wear affects how we support the introduction of interactive gloves into a generalized computing environment, and how we understand what the user experience is going to be like. Within the blocked-light context of VR, interactive gloves will mostly guide user interactions with virtual objects, with little expectation for interaction with the physical world around them. There may be challenges in donning and doffing the gloves, adjusting the headsets while wearing gloves (thin Velcro straps, moving the small IPD nub), or in using any other controller types in combination with the gloves, but generally VR applications will not require demanding physical object interaction.

This obviously changes in MR, where the user has full access to the surrounding environment. It is therefore reasonable to expect, as a first principle, that MR users should have equal access to virtual and physical objects. A challenge that arises, then, is that the glove used to provide access to the virtual objects may encumber their access to the physical objects.

Encumbrance covers a wide range of physical, mental, and social characteristics, but a basic definition emphasizes its nature to the user as an impediment or burden [58]. This reinforces that an encumbrance provides an unnecessary hindrance on a person's ability. Encumbrance can therefore be framed as a temporary disability – one that makes tasks harder that were previously easier to do. When the encumbrance is removed, the disability disappears.

1.1 Problem statement

Given this context, there is an opportunity to look at ways of measuring the degree of encumbrance that users might experience while using gloves to interact with common objects. For this study, the focus is set as the "Cost of Haptics" – given that it is a problem introduced by the addition of haptic devices to the hands, where haptics are used to create interactions focused on virtual objects. There is therefore an incurred cost to the user with the addition of haptic feedback, in exchange for a benefit in the form of enhanced immersion and other performance measures with the virtual object. It now seems critically important to quantify this cost, and to measure it against the perceived benefits.

The deliverable for this work is therefore an evaluation framework that provides a reliable and understandable measure for a few essential aspects of encumbrance. Creating this framework will require drawing testing and evaluation methods from an array of related fields – including industrial design, human-computer interaction, psychology, rehab medicine, and ergonomics. Each of these fields has developed their own approaches to evaluating encumbrance, and by drawing on their combined knowledge, it should be possible to structure an integrated framework that provides valuable metrics for designers. This set of metrics will therefore be available to inform the design of the haptic gloves and encourage the consideration of the cost of the encumbrance early in the design process while there is still time to make changes.

1.2 Smart Glove Design for MR

The design of smart gloves is a complex multidisciplinary endeavor, aggregating knowledge from many disparate fields. They are still novel devices, though there has been a long history of smart glove development – to date, none have been widely adopted. Yet smart gloves have still developed a common set of shared characteristics that define them.

1.2.1 Smart Glove Characteristics

Smart gloves can be distinguished from conventional gloves by the extended functionality they offer – usually because of a combination of these features:

1. Compute – Smart gloves are likely to have some version of on-board computing – either through the inclusion of a microcontroller or microprocessor.
2. Sensing – Most smart gloves will have a sensing capability that allows them to detect characteristics of the environment or wearer
3. Tracking – Many smart gloves will have tracking capabilities, which allow them to understand the position and trajectory of the wearer's hands
4. Actuating – Many smart gloves will also have an actuation function, commonly as part of a haptic display
5. Indicating – Some smart gloves also have indication capabilities, typically with light or sound, that can provide information to the wearer

The presence of any of these features differentiates smart gloves from basic gloves, as they enable functionality within an interactive system for their wearer. This greatly expands the functional scope of the gloves from barrier protection, thermal protection, and fashion to an increasingly broad range of interface possibilities.

Designing smart gloves adds a significant degree of complexity, as new constraints and characteristics start to dominate the user experience. Designers are now tasked with a much-expanded range of design considerations, each of which has a causal effect on the others. This entanglement makes even the simplest choice of where to place a component deeply complex.

1.3 Design Challenges

Part of the originating work that led to this dissertation involved the creation of the Haptic Mirror Therapy Gloves – a smart glove design exploration that provided early insight into the encumbrance effects of various design features. This project, discussed in greater depth in APPENDIX A, featured a series of divergent design explorations – each tackling a different aspect of the complex design process needed to create a glove to assist stroke survivors with their therapy.

Figure 1 – The Haptic Mirror Therapy Gloves.

The design exploration that underpinned the Haptic Mirror Therapy Glove project ultimately revealed some major design challenges with the creation of similar gloves. It became clear that smart gloves designs are inherently part of a complex functional and fabrication system, where every decision has a ripple effect on another aspect of the system. Each element therefore needs to be evaluated separately, and as part of the whole throughout the design process. This is particularly true when it comes to the proliferation of encumbrance effects linked to specific design alternatives.

One example of these encumbrance tradeoffs could be with the selection of conductor technology for the gloves – traditional materials and processes, such as wire and printed circuit boards (PCBs) provide excellent electrical performance but can provide a similarly high level of encumbrance. E-textiles and conductive inks can reduce the encumbrance of the design, but typically provide poor electrical performance. Flex PCBs provide a reasonable compromise. However, they are more expensive to produce, and challenging to integrate early in the exploratory phase of the project. Each one of these choices has tradeoffs and incur different ranges of encumbrance costs.

Another obvious example of encumbrance costs could be found with the selection of components to place on the fingertips. Fully featured smart gloves are likely to benefit from placing both sensing and haptics on the fleshy palmar side of the fingertip, as that is the primary area of control for dexterous object manipulation, and the site of the largest cluster of nerves on the finger. Placing components here helps create believable interactions, especially in VR, yet the components placed over fingertips directly interfere with physical object interactions, providing encumbrance that significantly affects the user experience in AR and MR. The fingertips are prime real estate for the glove designers, and deciding what components should be placed there, if any, presents a major challenge.

1.4 Ongoing design questions

These challenges uncovered by the Haptic Mirror Therapy Glove project lead to obvious questions – in such a complex design space, how does a designer know what the right choice for the selection and placement of components is? What tests can they run to help answer these questions? How are encumbrance effects best measured? Without an evidence-based approach, it would be difficult to gain useful answers to these questions.

It was obvious that there were encumbering consequences to each design decision, but there was no clear way to measure them. It was also apparent that the design of these gloves operated inside a complex system, where every decision had a ripple effect that influenced the choice around other designed components.

This design system was therefore full of tradeoffs – it is likely that good conductors would encumber fingers more, and that low-encumbrance conductors would prove to be unreliable. These explorations highlighted the relationship of fit to encumbrance and indicated is it likely that the tolerances for fit were much different for smart gloves than they were for more typical gloves.

Finally, the nature of the "Battle for the fingertip" became apparent. Both sensors and actuators need to be placed in the optimal position to interact correctly with the body – yet adding components to fingers encumbers their basic interaction with objects. It was clear there were costs to placing components on the body that needed to be accounted for, and new methods would need to be identified to better calculate those costs.

1.5 Research Goals

Under this broader umbrella of wearability evaluation, there was therefore an opportunity to identify a test regime that is capable of measuring encumbrance effects of gloves. The goal of this work was to therefore determine if these tests can show significant differences between gloves and be used to effectively evaluate and compare different glove designs. Given the determination of these differences, the studies would therefore allow the design team to identify the various costs of haptic glove features when interacting with physical objects. These costs, once identified, would yield opportunities to improve designs for haptic gloves and challenges that designers will encounter in creating haptic gloves for MR.

The studies performed in pursuit of this dissertation show that a mixed-methods protocol can be assembled and tested to show these differences, identify the cost to the user of their haptic gloves, and propose design solutions to address these challenges.

1.6 Structure of Research

Following this introduction, Chapter 2 provides background information on the prior art in measuring glove encumbrance, and various approaches to design performance evaluation. Chapter 3 details the start of a formal evaluation process, with Study 1: Glove Encumbrance Study. This first study evaluates a series of generalized hand functions tests and shows that significant differences can be identified between various gloves designs. This study also provides some detail into how the various encumbrance effects affect the user's experience and offers some initial guidance for how to mitigate these effects.

Chapter 4 details the modification and elaboration of the original protocol to become the Cost of Haptics protocol – which seeks to incorporate a refined set of tasks pulled from supporting literature, and create more accurate tests to better explore the effects identified in the first study. Chapter 5 then discusses the creation of a custom set of gloves, and the adaptation of the study to run during restrictions from the COVID-19 pandemic.

Chapter 6 details the data collection and results of Study 2: The Cost of Haptics Commercial Glove study. This study tests the same set of Gloves as Study 1, with the updated Cost of Haptics protocol, and confirms that the encumbrance effects are still detected, and shows the efficacy of the new protocol tests. Chapter 7 presents the findings of Study 3: The Cost of Haptics Constructed Glove study. This study tests a set of purpose-built gloves, each of which vary their designs by a single independent variable. This shows the Cost of Haptics protocol is sensitive enough to test small differences between the gloves and allows for the evaluation of a proposed solution for one of the dominant encumbrance effects.

Chapter 8 discusses the overall Cost of Haptics exhibited by the exemplar gloves that most closely resemble current VR glove designs. This includes a detailed discussion of the aggregate findings of the previous studies, a proposal for future work leading to an integrated framework, and a discussion of the value of this approach for designers.

CHAPTER 2. BACKGROUND

Evaluating the effects of glove encumbrance is not a new field of study – there has been substantial work on evaluating gloves in real-world scenarios, and these findings and methods are invaluable as a starting point for this study. The goal of this review is to therefore identify useful methods and metrics that can be used to evaluate the encumbrance effects exhibited by Mixed Reality gloves.

2.1 Encumbrance Evaluation

Iman Dianat has been studying glove effects on the hand for nearly a decade, and his comprehensive literature review – *Methodology for evaluating gloves in relation to the effects on hand performance capabilities: a literature review* – provides an excellent perspective on the current state of the field [19].

Dianat et al identify a range of possible encumbrances that gloves may impose on the hand, including manual dexterity, tactile sensitivity, handgrip strength, muscle activity and fatigue, physical discomfort, as well as effects on pinch strength, forearm torque strength, as well as finger and wrist movements. This list provides an excellent starting point for the framework, identifying numerous possible measures, and an overview of the methods used to examine them. The authors specifically state that they wanted to perform this literature review to "make recommendations for the testing and assessment of gloves" so their work provides an excellent precedent for the work examining glove encumbrance. Of particular interest is their review of the literature on manual dexterity, tactile sensitivity, fatigue, and discomfort, as those are the focus topics of this study, as they apply to smart gloves [19].

Manual dexterity can be measured using a variety of standardized tests, including peg boards and box and blocks tests, which allow for repeated measures of small object interactions [52]. These tests can be used to compare task completion times and error rates and these metrics can be used to infer the encumbrance effects. Gloves have been shown to negatively affect both completion time performance and rate of errors in a task [20,60]. There is also evidence to suggest that glove material and glove thickness may be a factor in aggravating these effects [59,84].

Tactile sensitivity is also affected by wearing gloves, as the glove material covers the receptors in the hands that would normally aid the completion of the task. Touch sensitivity is used to help detect pressure thresholds, spatial discrimination between points, and object identification, among other tasks [23]. To measure these aspects, common tools include aesthesiometers, such as the Semmes-Weinstein monofilament test, and the two-point discrimination test [7,10,85]. The results of studies using these tools are varied, as the glove material and variability in the application of the test seems to have an effect on the reliability of the measurement [19]. However, there is evidence to show that gloves affect tactile sensitivity negatively [89].

Comfort is another factor of glove use, and studies have captured subjective user feedback on this subject. Gloves may aid with comfort when they protect the user in object interactions that might be painful, but they may also provide perceived discomfort when the gloves are not serving that purpose [2]. Breathability was one aspect that affected users, as sweat accumulation can lead to discomfort, and will vary from glove to glove based on fit, pattern, and material [27,33]. Correct fit is another obvious contributor to encumbrance, as a mismatch in fit can lead to reduced mobility for the user, or an excess of material that may get in their way [88].

While this review from Dianat et al is comprehensive, and provides a strong base to build on for glove evaluation, it does not consider the factors that affect interactive gloves for VR and MR. The gloves reviewed were all conventional gloves, typically used in industrial contexts, and did not have smart glove design features such as actuators or integrated circuits. There is a much larger body of work for VR gloves, as compared to MR gloves, as the researchers are typically concerned with the ability of the VR gloves to render haptic experiences, than with their interactions with the physical world. However, many VR studies repurpose tasks like those used in Dianat's review, so it may be possible to use these same methods for the evaluation of smart gloves.

One precedent that attempts to compile a comprehensive set of tests is Catoire et al's *Towards a Test Battery to Benchmark Dexterous Performance in Teleoperated Systems* [15]. The goal of this paper was to identify a set of tasks that could effectively and comprehensively benchmark system dexterity for the remote operation of robots. As this need represents a wide range of tasks and operating contexts, their team determined the need to test a broadly generalized set of dexterous tasks. This is an ambitious challenge, as it anticipates unspecified future use-cases, and attempts to test a wide enough range of metrics to accommodate them. MR gloves present a similar set of challenging characteristics, so the choices Catoire et al made in the construction of their test battery should be generally instructive.

This paper details their evaluation and selection criteria. They ultimately chose five diverse tasks – many of which were discussed in Dianat's review. The task list selected included the Box & Blocks test [52], the Purdue Pegboard test [83], the Minnesota Manual Dexterity test [18], the ISO 9283 trajectory test [38], and adapted version of the IROS 2017 screwing sub-test [37]. The first three tests on that list show a range of object interactions at three different scales: 1" cubes, thin short pins, and 1.5" wide disks. The cited ISO 9283 test is a tracing test, and the IROS 2017 test involves connecting two objects together using screw threads. These tasks feature tool manipulation, dexterous object manipulation, object grasping and release, and tests of speed and accuracy. Collectively, the authors of this paper believe that these tests provided them with a first step towards evaluating hand function for generalized dexterous use-cases. It seems likely that this approach will work well for MR smart gloves as well.

Many studies used task completion time as a performance metric, including those that asked participants to sort objects by relative perceived weight and scale, to position virtual objects, and to play a virtual piano keyboard [34,35,57]. Error rates were also used to show performance effects, as well as accuracy [73,86]. Additionally motion-tracked VR studies were also able to gather more detailed motion data from their system sensors, including range-of-motion, arm trajectory, and hand velocity – this is a key benefit of VR studies, as this type of data are otherwise very challenging to collect [39,49].

There are a very small number of studies that have looked at hand performance in interaction with physical and virtual objects. Magdalon et al used a Cyberglove to test object interaction in virtual environments and physical environments and noticed that interactions wearing the gloves (in both VR and real life) were slower and less accurate than the same tasks performed without gloves. This yielded a useful data from the VR sensor system and provides a great precedent to build on, as it provides a model for future study of glove encumbrance with trajectory data.

2.2 Wearability Evaluation

One of the first consideration that designers face with wearable technology is proxemics – the human perception of space. Gemperle argues that there is a tacitly understood envelope around the human body where we can add volume without disrupting the perception of the whole form as human [24]. This volumetric space is largely centred around the torso, which is why we can wear large backpacks without looking strange to a viewer. That same backpack shaped volume placed on our feet may look comical or unsettling. This consideration is particularly challenging for the design of smart gloves, as the extremities are more sensitive to this phenomenon – hands can have very little volume added before they start to trigger this effect. The prospect of looking abnormal while wearing technology on your body also raises concerns around social acceptance. It is still relatively uncommon to see strangers interacting with smart gloves, and may seem unnatural to viewers in public space [70].

This limits the opportunity for form massing in the design, as does the increased constraints on weight distribution. Adding weight to the hands in relatively small amounts can provide a significant encumbrance to the user [43]. Body mechanics and reachability are also factors, as there are limited ranges of movements that are comfortable over an extended duration. Placing components on the body requires careful consideration to make sure they don't limit movement, or access to the interface [91].

Thermal tolerances are also vital to consider – gloves already provide a thermal envelope around the hand but adding compute elements to the glove increases this thermal load. Managing sweat becomes a significant consideration for glove design, especially those that support high intensity activities like gaming. Other considerations include biometric sensing and haptics, which need precise placement on the body to interface with the correct body part [26]. Additionally, there are ongoing concerns about the manufacturability of these garments, as they require new processes, many of which have yet to be developed.

2.3 Design Performance Evaluation

Design projects contain inherent ambiguity in selecting their optimal path to completion, therefore the adoption of a structured design process with feedback has long been given as a solution to manage complexity. Many design process models have been offered to structure this iterative exploration – some of which couple interaction design process with evidence-based design process – which attempt to bridge the gap between traditional design, and the scientific method.

The most basic illustration of design process may be Damien Newman's Design Squiggle [61]. This is not a model that a team can build a detailed project plan around, but it shows the most essential path of progress for a design team – moving from uncertainty of what to build, to clarity of what to build. The way to progress from one side of the squiggle to the other is to make models – artifacts whose creations helps the designer to understand whether they are building the right thing.

These models have a role to play at the beginning of the process, at the point of greatest uncertainty. Many designers reserve the word prototype as a special model – one that is prototypical of the final design – and the last model that designers are able to make changes to before the production process begins. At the end of the squiggle is the design itself – the resolved pattern that describes the product that will be manufactured and replicated. The line imposed on top of the squiggle shows the cost of making changes at any given stage. This is inherent in a profession that makes use of expensive injection-molded tools and validation testing production runs, that the cost of changes will continuously increase as the project develops.

The second influential design model is the Double Diamond – popularized by the UK Design Council [16]. This is the primary iterative model taught to designers, guiding them through periods of divergent thinking (taking in new information) and convergent thinking (refining and eliminating information). This process creates feedback loops and inflection points, which measure the success of the design effort against the project brief. An evidence-based approach converts the diamonds into tests of a hypothesis concerning the fit of a given project solution. As with the squiggle, designers move through this process by making models, and then evaluating the success of what they have built.

The models built during these processes have typically followed the traditional pattern of form and function models. These may include sketching, foam models, and appearance models on the form side, and schematics, breadboard circuits, and printed circuit boards on the function side. These approaches are valuable, but they do not directly address the design of the interaction itself, which is vital to the success of the project. Here, there is an opportunity to introduce an "action model" that incorporates Interaction Design prototyping techniques - such as bodystorming, paper prototypes, and storyboards. The goal of the action model is to directly design and test the interaction and use it as part of the feedback loop to match the form and function.

This produces a set of "looks-like", "works-like", and "acts-like" models that are each relatively easy to develop and refine using their own techniques, but that collectively inform the overall design. Keeping these modelling processes separate – at least at first – allows for rapid iteration at a lower cost. However, the findings gathered at from each new model to influence the next in the series, and the models in the adjacent paths. Eventually, as the designer starts to understand the relationship of all the design elements, they work over the project to integrate these separate modeling processes into a single integrated prototype.

The integrated modelling process therefore takes inspiration from the Design Squiggle and Double Diamond design process to take advantage of the three distinct modelling processes, and to integrate them into a single system – as shown in Figure 2. This process model starts with each of form, function, and action models separated, with the goal of eventually integrating them. To do this, the designer starts with one, develops it, and creates the next model type with information gained from the first. This continues through a period of divergent thought until the designer feels they have enough information to test their design hypothesis, at which point they move to converge and integrate the three models together.

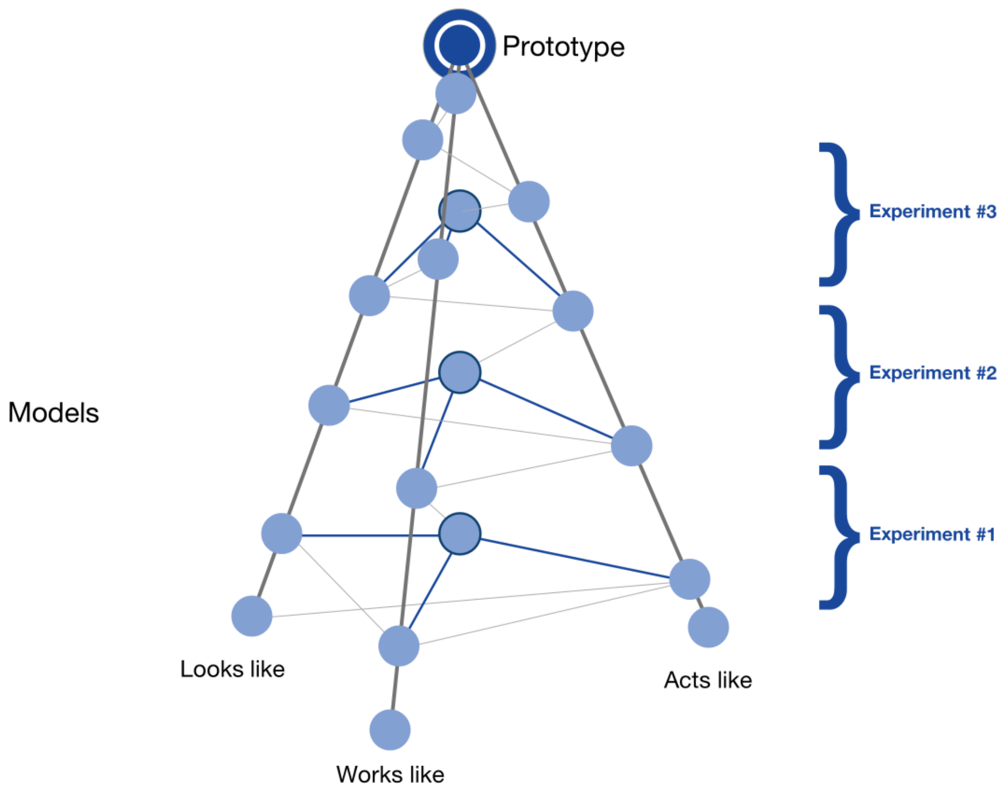


Figure 2 – The integrated modelling process.

The resulting integrated model allows them to measure their success against the requirements for the project, and to make plans for the next stage of the design process.

After evaluating the integrated model, they return to the separate modelling streams, with the intent of integrating them again when they have gathered enough new information. Returning to the separate modelling processes allows for more rapid iteration and exploration, while converging to an integrated model encourages repeated testing. In this way, the designer begins to move through the uncertainty of the design process and converges on the end goal of the resolved prototype.

2.4 Design Iteration and Longevity

The application of iterative design models differs between design disciplines, each of which are driven by differing trend cycles, market forces, and the durability of their goods. These factors can play a significant role in determining how iteration and feedback is managed, and how risk tolerant the design team may be in the development of novel products. Placing wearable technology – particularly smart gloves – along this continuum can illustrate where the evaluation challenges and opportunities might be found.

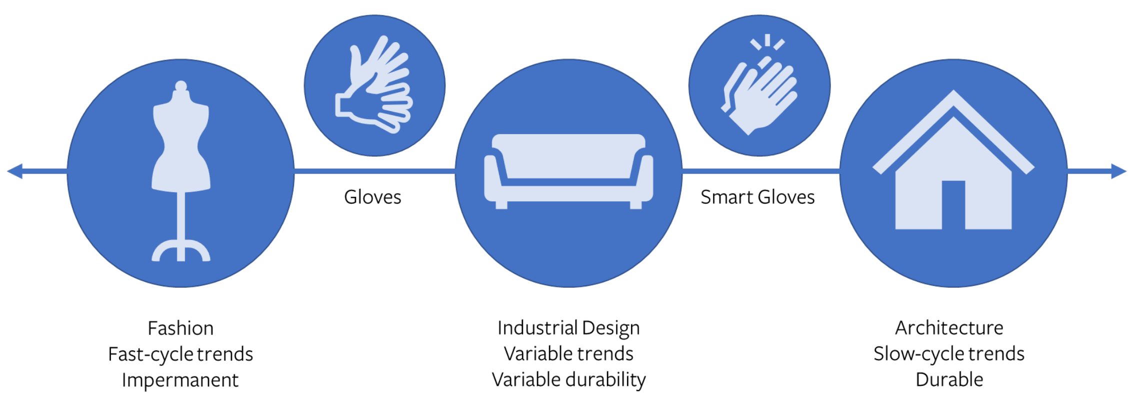


Figure 3 – The design iteration and longevity continuum.

Fashion anchors one end of this spectrum, as seen in Figure 3, as it is motivated by fast-cycle trends, and its designs are largely impermanent. With such rapid turnover in their product lines – often a complete refresh every season – fashion designers have had to adapt to rapid iteration. This is suitable to their materials and design process, however, as deep insight can come from straightforward activities, such as draping fabric on a dress form. The feedback loop for fashion iteration is very short and direct and offers little risk to experimenting with new ideas.

This is contrasted with architecture, which anchors the other end of the iterative spectrum. Architecture trends are comparatively slow-cycle, and their products are very durable. Architectural iteration happens largely at the modelling stage, as it is cost prohibitive to build out a full-scale building to test new iterations. Where fashion designers can run multiple versions of a garment in a single week, architectural designers are forced to wait for construction to begin to gain any hands-on experience with their product. For this reason, architecture and building construction have been eager to embrace detailed simulations that try to capture as much detail of the building as possible – this is the canvas for their iteration, as it allows them to make changes and test the outcomes earlier in the design cycle where the risk is reduced.

A third discipline – Industrial Design (ID) – straddles the middle ground between the prior two. ID can vary how it responds to trend cycles depending on the project and can similarly vary the durability of their projects. This means that a broad range of design processes can be engaged, depending on the needs and characteristics of the project – and with this variance follows the iterative mechanisms that are available. For projects where the product is mass-manufactured, based on existing lower-risk processes with known materials, the iterative model-making approach is more than sufficient to answer the questions required by form development. More complex projects benefit from more methodical and research-oriented processes as the risk increases. Over time, products may move from one end of the scale to the other – as an example, lower-cost mobile devices have become commoditized and are now therefore easier to develop and assemble.

Placing smart gloves on this continuum should therefore help illustrate some opportunities to fit the more suitable iterative strategy to the problems faced by their designers. Conventional gloves fall on the fashion side of the scale, as there is little risk in trying out new patterns and materials, with feedback coming quickly. Smart gloves, however, are anchored clearly on the architectural side of the scale, due to their complexity, novel processes, and their risk in bringing them to market.

The simulation and research approach to iteration taken by architecture may therefore provide benefits to the smart glove design process. Simulation presents a unique challenge here, however, as the humans in architectural systems are modeled in aggregate – looking at the group behaviors and outputs that drive decisions about resource allocation and placement. For smart gloves, a simulation of the glove itself would be challenging, but would be simple compared to simulating the human wearer of the glove – with all their inherent cognitive, behavioral, and physiological characteristics. This suggests that an evidence-based design process, based on human-centered design principles, with a regular cadence of human subject research may offer the best structure to iterate and evaluate through the challenges smart gloves face on their path to market.

2.5 Incorporating the Scientific Method

Following this conclusion, smart glove projects should benefit from the introduction of evidence-based design processes. One of great benefits of introducing a scientific approach into the research is the quality and variety of evidence that the designer can access. This leads to new challenges, as there are two distinct stages to incorporating evidence – the evaluation of the evidence itself, and the incorporation of it into the design [66]. The former builds heavily on the existing scientific method, while the latter draws on the traditional design skills. It is ultimately still up to the designer to respond to how the evidence informs their work.

While there are many possible ways of gathering evidence that benefits design research, one of the most directly applicable methods is the small-N study, so-called because it they are built around a small number of participants. Small-N studies work well with design research as they are not seeing causal evidence – rather they are designed to prove the presence or absence of an effect [28]. For designers, these provide structured ways of testing design alternatives, by constructing models with well-controlled variations in the design, and then testing them with users. If the models are properly controlled, they may exhibit an effect of some kind (i.e. performance metrics) which can be measured and used to compare the two variations. This dovetails well with performance-based design approaches used in building construction, and provides an accessible framework to assess quantifiable data that help make design choices [4]. While a single A-B test can show the presence of effect, running a replication study (i.e. A-B-A-B) will improve the validity of the results.

Designers are by their nature more familiar with qualitative methods, due to the integration of design ethnography and observation tools into design practice. However, qualitative tools can be selected that complement the Small-N study design. These include post-task questions, scaled response questions, and semi-structured interviews. The benefit of a post-task question is the ability to gather and track data at every experimental condition of the study – and there are many tools available to structure these questions, such as the NASA-TLX scale and the Single Ease Question (SEQ) [31,75]. Scaled-response questions and semi-structured interview tend to work well together – especially if the pilot yields a number of similar responses in the interview that can be translated into a scaled-response question [76]. Assembling this collection of qualitative and quantitative data therefore gives the designer tools to prove out the actual performance of their designs.

2.6 Design Performance Testing

Precedents for design performance testing strategies can be found again in building construction literature – a design discipline that thrives on evaluating performance of designs well in advance of their construction. Godfried Augenbroe has authored a paper titled *The role of simulation in performance based building*, which lays out a strategy for using key performance indicators to evaluate design alternatives [4]. While the focus of the paper is on applying this strategy to building construction, his methods can be translated to any other design discipline.

To structure the performance evaluation system, Augenbroe identifies two facing sides – one of the system requirements, where the user's needs are identified, and the other with the technical systems that attempt to meet those needs. Each side starts with generalized categories, and branches down into increasing detail, as specific requirements and sub-systems are identified. The point where the two sides meet are performance indicators – single discrete measures of quantifiable elements of the technical system whose change shows a degree of efficacy in meeting the requirements. Augenbroe bases his analysis around identifying these performance indicators, and then monitoring their changing values across varied designs in the simulation. As the indicators move, the relative performance of the different design alternatives is shown – guiding designers to make educated choices and improvements.

As discussed earlier, smart glove simulations are challenging to envisage – due to the human element of the system – but Augenbroe's approach can still translate if human subjects research is substituted for the simulation. In this case, performance indicators can be identified that stem from quantitative measures (or well-structured qualitative measures) gathered over the course of the study. If these measures are shown to vary solely based on the glove condition being tested, they can be used to infer the performance of the various glove designs. As the encumbrance measures reviewed typically provide clear quantitative measures – speed, accuracy, sensation thresholds – these are good candidates for use in a system like the one that Augenbroe describes.

A further benefit to this performance evaluation system is the means of displaying the performance results visually. Augenbroe makes use of radar plots to collect and arrange the performance indicators in a single chart, which allows the multivariate data to be plotted in a way that easily shows differences between two test conditions. The performance indicators are normalized before they are arranged in their scales, which allows the viewer to quickly infer which design perform best across the entire range of indicators. This system allows designers to make quick comparisons of the similarity between designs, as the general shape of the plot communicates their strengths and weaknesses.

The structure provided by this performance evaluation strategy should make it relatively easy for designers to identify the metrics that will best help them make choices, and to specify studies that present a high degree of value for the design process. These studies can be generative – testing small design features early in the cycle that allow quick evaluation and iteration – or summative, by establishing a final benchmark measure at the end of the design cycle. This system should help designers better grapple with the complexity that smart gloves present, by giving them better feedback, sooner.

2.7 Research Opportunity

There was therefore an opportunity to draw from the ergonomics and building construction literature to design a system to measure glove encumbrances for Mixed Reality, following a mixed-methods, task performance approach, with the goal of evaluating design alternatives. The following chapter will discuss the initial exploratory work that was completed looking at the design features intrinsic to smart gloves, and the initial observation of their effects.

CHAPTER 3. STUDY 1: GLOVE ENCUMBRANCE STUDY

3.1 Introduction

This study was created to look at ways of measuring the degree of encumbrance that users might experience while using gloves to interact with common objects. The goal was to have users perform basic tasks while wearing various gloves, and to record both quantitative performance data, and qualitative experience data to better understand how the encumbrance had affected them. The project was designed with MR in mind, with the intent that there might be opportunities to improve the user experience in VR as well.

3.1.1 Glove Designs

The study focused on identifying a set of commercially available gloves that had design characteristics that might be considered in future MR-glove projects. This allowed a broad evaluation of different types of encumbrance across a range of materials, patterns, and construction methods.

After purchasing and reviewing a broad set of designs, a final set of gloves was selected which could be arranged along a general continuum of encumbrance, as seen in Figure 4.

Figure 4 – 6 stages of glove encumbrance.

3.1.2 Powermesh glove

These gloves are commonly used for motion capture applications. They are very light-weight – constructed out of stretchy but smooth knit open mesh – and conform easily to the hand. These were the only gloves that allowed direct access to the skin of the hand through the material.

3.1.3 Running glove

These are light-weight thermal gloves and are commonly used as glove liners or for short jogs in cool weather. They are made of a stretchy 55% polyester/45% nylon with a smooth finish. These gloves have conductive fabric on the index finger and thumbs to allow for touchscreen interaction.

3.1.4 Gardening glove

These were medium weight gloves, used for protection from sharp plants while gardening. They are made from a knit Nylon – created in a single process with no seams on a machine – and then dipped in Nitrile to cover and protect the palm. They are breathable on the back, and still somewhat stretchy.

3.1.5 Tactical glove

These are heavier weight and used for outdoor activities like paintball. They are made from a suede-finish Nylon, with thermoplastic rubber armor added to the back of the hand and fingers. The gloves are not stretchy, and are stiffer to move in, with larger than average fingers.

3.1.6 Hockey glove

The heaviest of the gloves tested, these are used for junior ice hockey and road hockey. They are made of a variety of materials, with a thin synthetic leather palm and a sport-knit outer layer that contains thick foam padding. The fingers are curled into a natural state for holding a stick, and the cuff is quite loose.

Having assembled and tested each of these gloves, a study was proposed to see how participants would be able to use these gloves to complete common tasks.

3.2 Apparatus

The tasks for this study were selected to focus on object interaction at the fingertips. Inspired by some of the testing protocols used by Occupational Therapists to evaluate hand function, a set of six tasks that tested a broad range of general hand function were selected.

3.2.1 Typing Test

Typing is a common baseline task, with well understood performance metrics. It is a useful task to measure the encumbrance effects imposed by the gloves on a well-developed skill that each of the participants already had. Typing is a common skill, requiring speed, dexterity, accuracy, and precision, and most people have invested some time in learning this skill.

In this task, participants completed a 60-second online typing test with each glove. After each glove task they were asked for their evaluation of the activity, provided a NASA-TLX score, and recorded their WPM and error rate scores. Following feedback from the pilot, scaled-response questions related to comfort, typing experience, and eye position were asked.

3.2.2 Monofilament test

To measure hand sensation, the Semmes-Weinstein monofilament test was selected, which can measure the force detected by a participant against their fingertips down to fractions of a gram. The test uses calibrated filaments, which are pressed against a participant's fingertips until they bend at a 90-degree angle, at which point they deliver the rated force. Participants are blinded and asked to indicate verbally whenever they feel a force on their fingers – the smallest force they can successfully detect three times in a row is then recorded as their score.

3.2.3 Box and blocks test

The box and blocks test is also frequently used by Occupational Therapists to measure hand function and response time. The apparatus included a double-sided box with a low barrier across the middle, and 150 blocks in 4 different colors placed on one side. A modified version of the protocol was implemented, as it was deemed to be better suited for participants without severe disability. This modification was asked participants to sort the blocks by color into two sets and move all blocks of one set to the side of the box. Participants were asked to go as fast as they could, and their times were recorded.

3.2.4 Maze tracing test

In this task, participants were asked to trace the solution to a maze. The solution was indicated with a contrasting color along the route, so participants did not have to work to solve the maze – only to trace the pen through it. Participants were asked to go as fast as they could without hitting the walls of the maze with their pen, and their times and errors were recorded.

3.2.5 Mr. Potato Head test

This test used a Mr. Potato Head toy to create an assembly task using different shapes. Participants were provided a visual reference showing a Mr. Potato Head to build out of various parts and were asked to replicate the character by drawing parts out of a bin. Participants completed the task as quickly as they could, and their times were recorded.

3.2.6 Geometric solids test

This task asked participants to find specific objects in a box of similar shapes while blinded. Participants were shown a picture of an object, asked to close their eyes, and then asked to find the object in the box as fast as possible, using only their sense of touch. The test was repeated three times with each glove, and the participant's times and errors were recorded.

3.3 Methods

3.3.1 Participants

Participants for this study were recruited from the general population, with a sample of 12 adult participants. The protocol was approved by Western IRB, and the Georgia Tech IRB (in reliance on the Western IRB approval), and all participants provided their written informed consent.

3.3.2 Study Design

This study was structured to test each combination of glove condition and task, while controlling for ordering effects. This was done by using a latin square to structure the order of the glove conditions and using a random generator to structure the order of the tasks within each glove condition. Six glove conditions are included in this study, which are discussed in detail in Chapter 6:

- Glove 0: no glove
- Glove variant 1: User Study Glove - Powermesh
- Glove variant 2: Running glove
- Glove variant 3: Gardening glove
- Glove variant 4: Tactical glove
- Glove variant 5: Hockey glove

Following the welcome and consent participants were given their first pair of gloves, and then asked to complete the monofilament test. They would then be guided through the four other tasks (in random order), recording the data as they were completed. After each task, the participant was asked to complete the Single Ease Question (SEQ), and then asked to describe their experience in a brief semi-structured interview. Following the conclusion of all the tasks for a single glove, the participant was asked to complete a set of scaled response questions.

This pattern was followed for all five gloves, and a sixth baseline condition without a glove, which was included in the random order. Participants were given a break after 3 gloves were completed, where they were asked to scan their hands on a flatbed scanner. Following the conclusion of all six gloves, participants were asked to take four measurements from their hands and were provided a set of summary scaled-response questions to answer. The entire session took around 3 hours to complete, and was recorded on a video camera, with extensive session notes captured by the facilitator.

This protocol was designed as a mixed-methods study, intended to look for triangulation between quantitative performance metrics, and qualitative experience metrics. This approach supported the study to identify the broader effects of encumbrance, some indication of the causal reasons for the effects, and to offer some design recommendations to address them.

3.3.3 Measures

Each of the five tasks had a primary measure of performance, and a set of associated subjective measures. The performance measures are as follows:

Table 1 – Performance measures for Study 3 tasks

Task	Measure	Unit	Calculation
Typing test	WPM	count	Correct, complete words / minute
	Error rate	%	Correct words / total words
Monofilament test	Force	g	Smallest force detected
Box and blocks test	Speed	seconds	Time to completion
Maze tracing test	Speed	seconds	Time to completion
Mr. Potato Head test	Speed	seconds	Time to completion
Geometric solids test	Speed	seconds	Time to completion

The subjective measures for this study are discussed in depth in Chapter 4.3. The set included for the study are as follows:

- Single Ease Question (SEQ) [1-7]
- Scaled response typing questions [1-5]
- Scaled response glove experience questions [1-5]
- Glove preference questions [choice of preferred glove]

3.3.4 Analysis

Analysis for this study was completed using SPSS for Windows. Performance data for these gloves and scaled response questions are considered to be non-parametric, therefore the Kruskal-Wallace test was selected to determine if there was variance between the median values for each group. The results are rendered in boxplots, which depict the median, quartiles, and outliers. SPSS renders the box as representing the interquartile range (IQ) of the central 50% of the values (ranging from the 1^{st} to 3^{rd} quartile), with the median indicated as a line across the box. The whisker lines show values that extend from the limits of the box to the highest and lowest values that do not exceed 1.5 times the range of the IQ. Circles and crosses indicates outliers, the latter of which represent extreme outliers with values that exceed 3 times the range of the IQ [42].

3.4 Results

3.4.1 Typing test

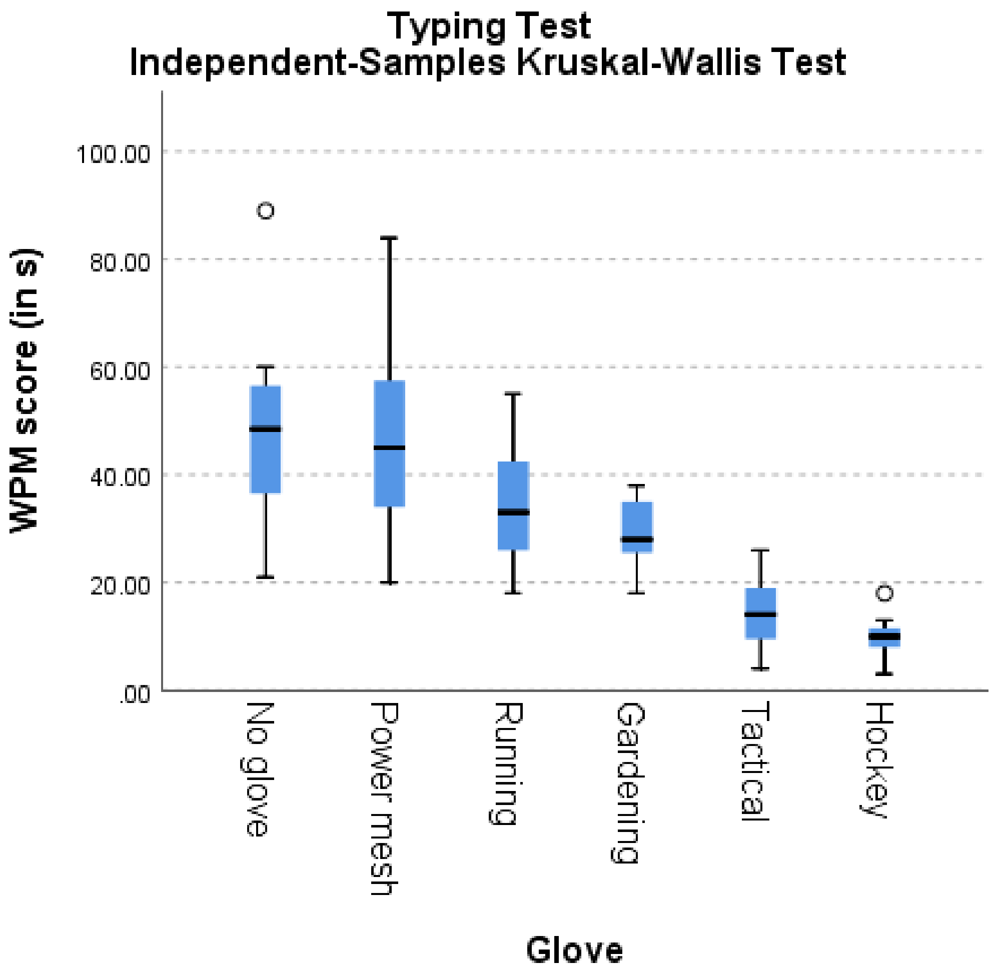


Figure 5 – Typing test Kruskal-Wallis test results – there is a significant difference in the WPM score between gloves, $\chi2(5) = 51.10$, $p = 0.000$

The results for the typing test showed a clear difference between typing with no gloves, lighter weight gloves, and heavier gloves. Figure 5 shows the difference in WPM scores between each glove, ranging between the median "No glove" score of ~50 WPM, and the Hockey glove median score at ~10 WPM. The median WPM score consistently falls as the relative encumbrance of the gloves rises, with the hockey glove showing low variability and consistent results across the entire population.

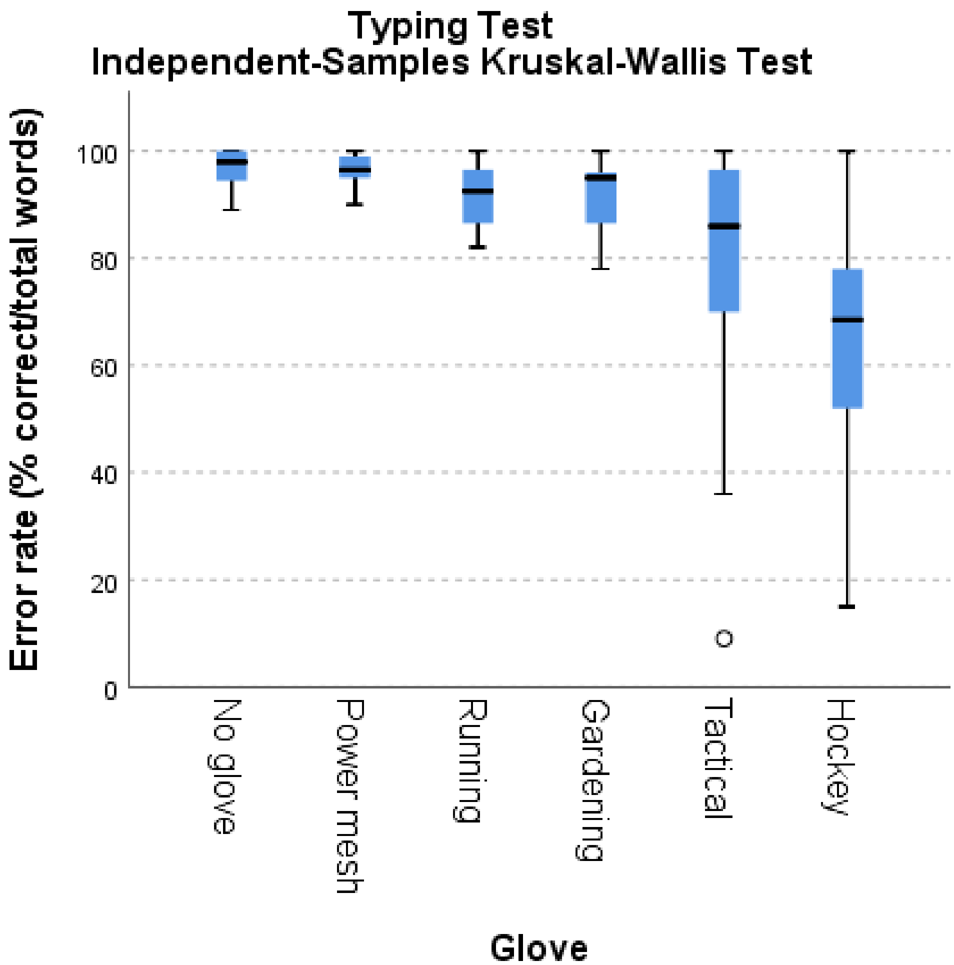


Figure 6 – Typing test error rate Kruskal-Wallis test results – there is a significant difference in the error rate between gloves, $\chi2(5) = 27.12$, $p = 0.000$

A similar chart shows the difference in typing error rate between gloves is seen in Figure 6. The hockey gloves showed a range in error rate from ~15% to 100% – this wide variability is likely attributable to the task adaptation participants undertook in response to the gloves. Participants were observed to slow down their typing speed to focus on carefully completing a few perfect words, or were so frustrated that they were unable to complete any words at all.

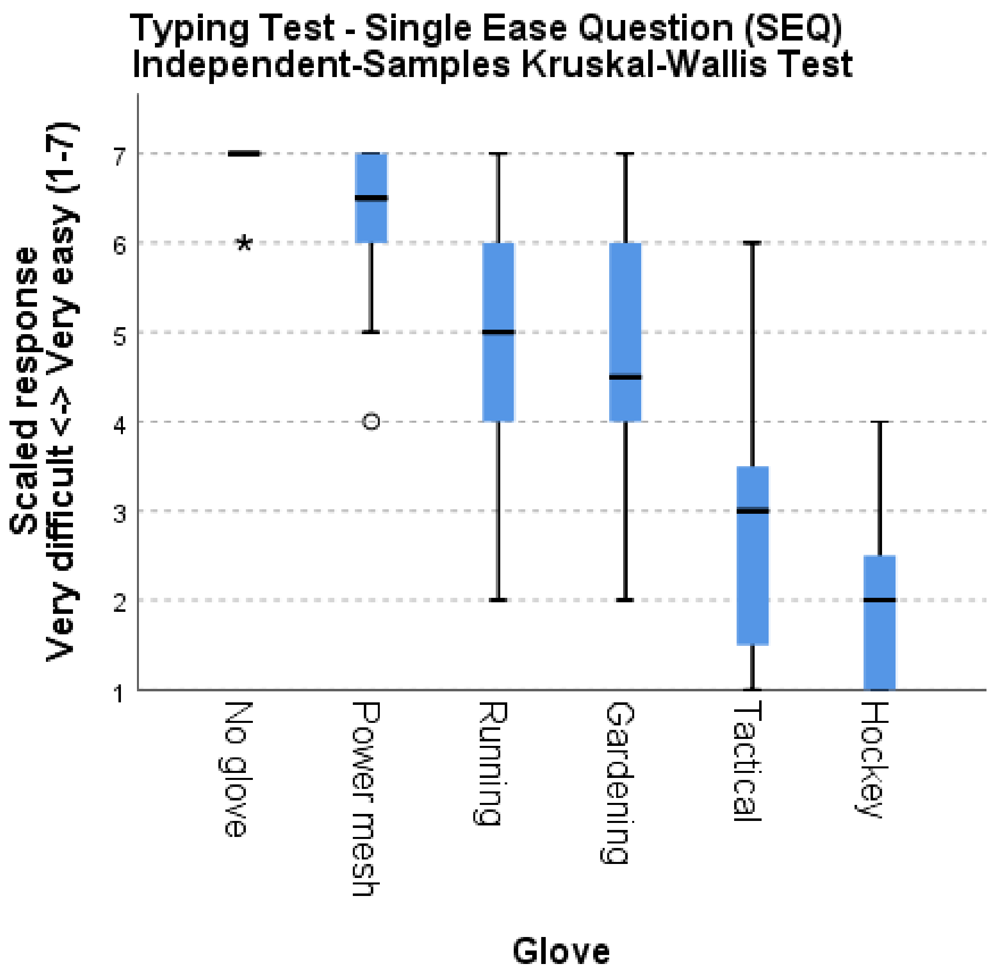


Figure 7 – Typing test Single Ease Question Kruskal-Wallis test results – there is a significant difference in the perceived difficulty of the task between gloves, $\chi2(5) = 15.95$, $p = 0.007$

The SEQ response questions tells a similar story – participants saw noticeable changes in the difficulty of the task, and its effect on their typing ability as they progressed through the gloves. Participants reported all the way through the test that they were having to think more about the task than they had expected, especially as the gloves got heavier. This suggests that the cognitive load of the tasks increased with the weight of the glove, as participants struggled to figure out how to adapt their typing style to work around the glove encumbrance. These effects are shown in more detail in the scaled response results in Chapter 3.5.

Overall, the gloves made from thinner material performed better. As the material over their fingers got thicker, participants reported a loss of sensation, as they could no longer differentiate the edges of the keys from each other. Eventually many participants lost the ability to tell keys apart, and some reported they could no longer tell when keys were being pressed. This effect seemed to be more pronounced on touch typists, who had to work harder to compensate when their skill was disrupted. Participants verbally expressed a strong preference for the Power mesh and Running gloves, and the scores across each of the metrics validated this preference.

3.4.2 Monofilament test

Table 2 – Hypothesis test summary for Kruskal-Wallis test across all three Monofilament test conditions. Significance level is 0.05, total N=72.

Task Condition	Test Stat.	Sig.
Monofilament test - Thumb	45.96	.000
Monofilament test - Index	47.28	.000
Monofilament test - Pinky	43.40	.000

The monofilament test showed clear differences between the gloves, as shown with the clearly significant results across all tested fingers in Table 2.

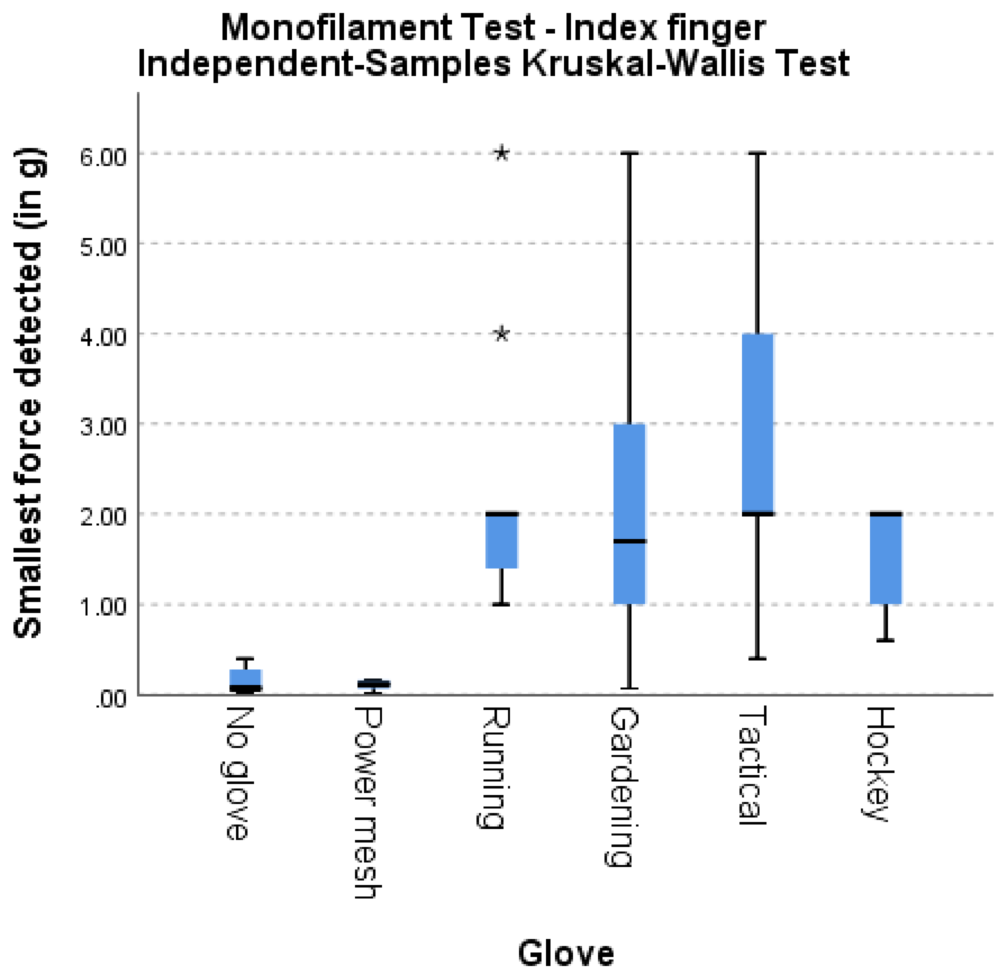


Figure 8 – Monofilament test Kruskal-Wallis test results – there is a significant difference in the smallest force detected on the index finger between gloves, $\chi2(5) = 47.28$, $p = 0.000$

To look at the representative example of the Index finger results, as seen in Figure 8 – it is immediately obvious how much the thickness of the glove material affects the participant's ability to detect small forces. The typical participant was able to detect a 0.16g force with their bare hands and the Powermesh glove, but the median force detected for the remaining four gloves was close to 2g, with a high degree of variability. This is likely due in part to the gloves that had doubled material on the thumb for reinforcement – the running glove, for example, has the touchscreen fabric on the thumb and index finger, which was responsible for doubling or quadrupling the smallest detectable force, as compared to the pinky which had a single layer of fabric. Of note is the relative success of the hockey glove, as the bulk of the glove is situated on the dorsum of the hand. This contrasts with the poor performance of the tactical glove, which has thick pads over the fingers. This test quickly demonstrated the impact that a single piece of fabric can have on participant sensation and perception and demonstrated a low-cost way of measuring this type of encumbrance in future studies.

The task performance results are mirrored in the subjective SEQ results, seen in Figure 9, which show a clear change in the perceived difficult of the task from a median score of 6 ("easy") to the hockey glove's median score of 3 ("somewhat difficult").

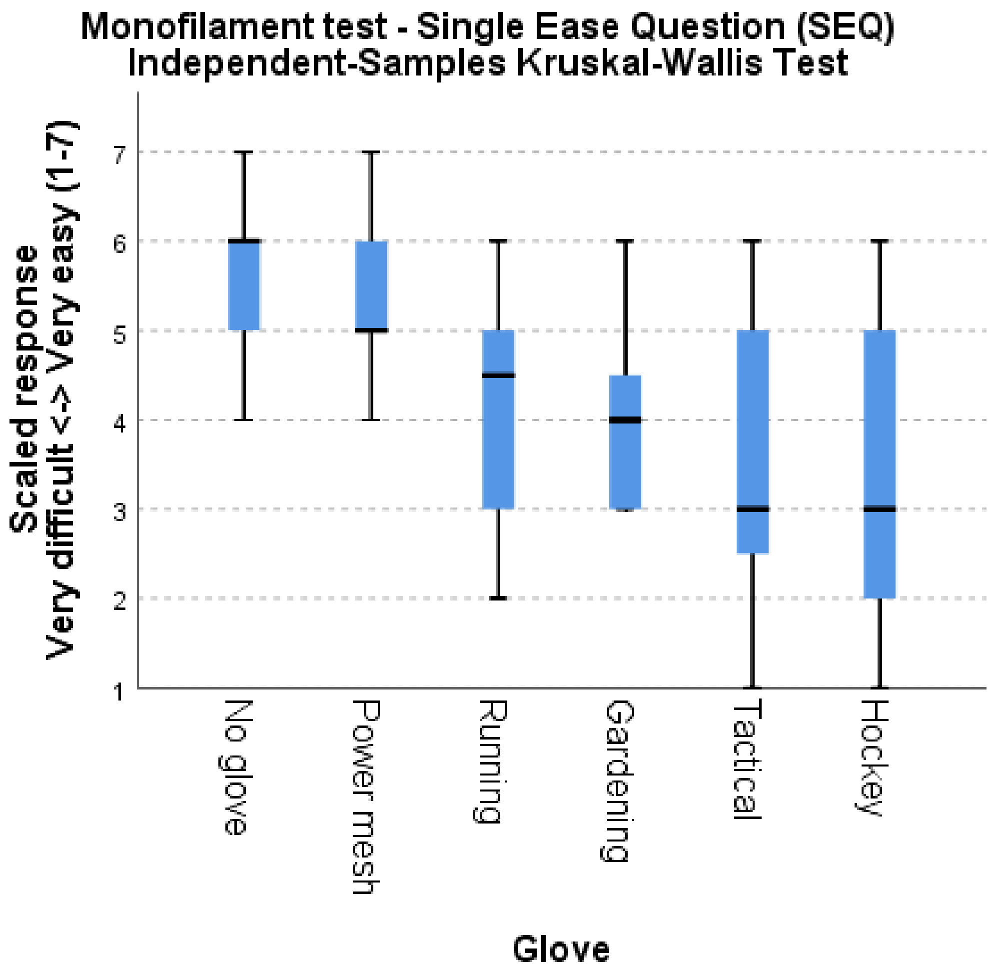


Figure 9 – Monofilament test Single Ease Question Kruskal-Wallis test results – there is a significant difference in the perceived difficulty of the task between gloves, $\chi2(5) = 25.95$, $p = 0.000$

3.4.3 Box and blocks test

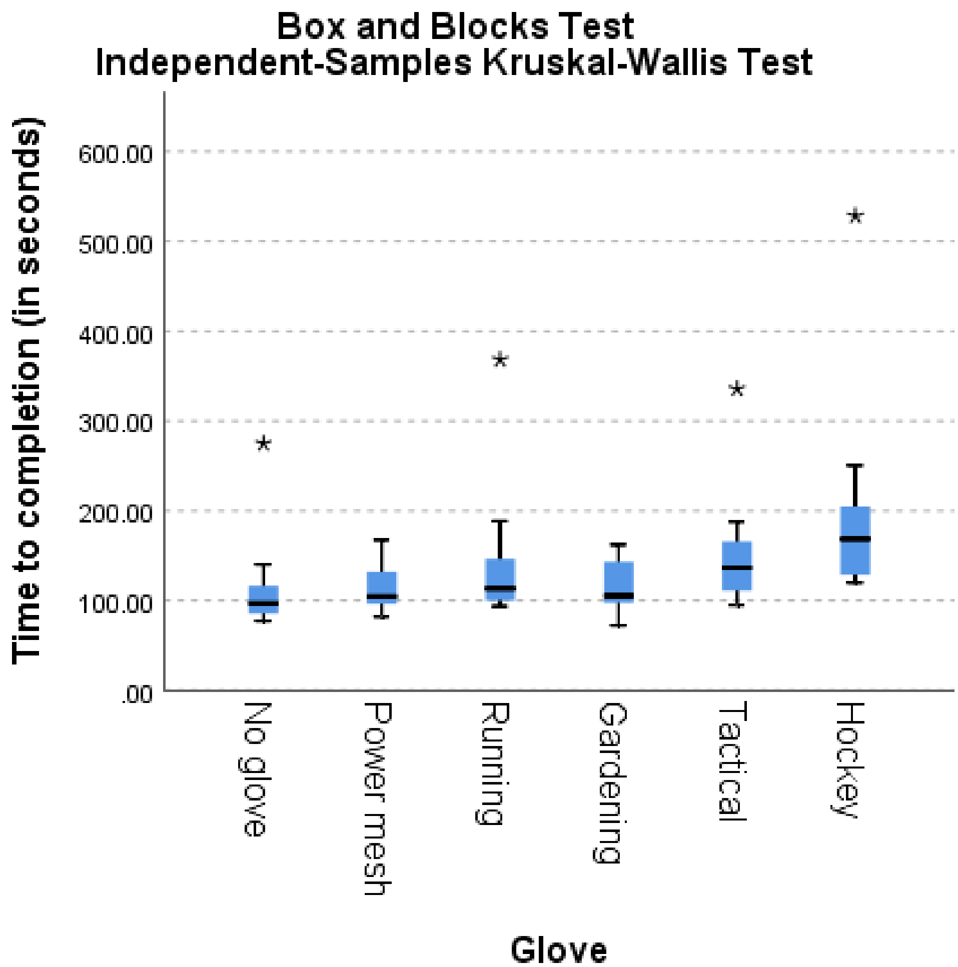


Figure 10 – Box and blocks test Kruskal-Wallis test results – there is a significant difference in the time to completion score between gloves, $\chi 2(5) = 21.16$, $p = 0.001$

There were clear differences in the median times between the gloves in the box and blocks test – generally following a predictable pattern of increasing the times as the weight of the gloves increased – notable in Figure 10. There was one exception – the gardening gloves slightly outperformed expectations and yielded a similar median time to the Powermesh gloves. As Figure 11 show, participants reported a strong preference for the increased grip the gardening gloves provided, and this showed up both in their improved times, and in the preferential ratings with the SEQ score.

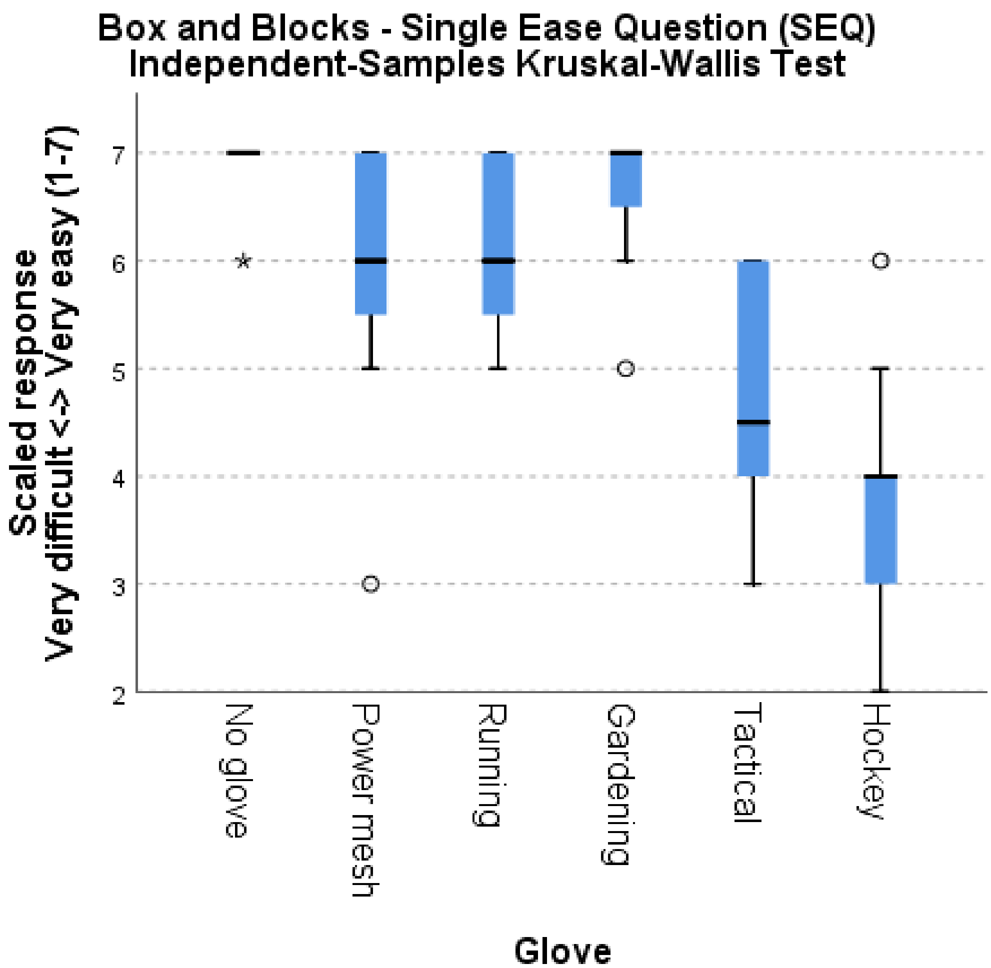


Figure 11 – Box and blocks test Single Ease Question Kruskal-Wallis test results – there is a significant difference in the perceived difficulty of the task between gloves, $\chi 2(5) = 44.20$, p = 0.000

Overall, participants reported that higher encumbrance meant a greater cognitive load. In this case, like the typing test, that meant they had to spend more time and effort thinking about the task, rather than doing it automatically. One of the most notable changes was visible with the grip participants used to pick up the blocks – with their bare hands and the gardening gloves there was a high likelihood they could use many different grips interchangeably, without extra effort. With the other gloves, most had to adapt their tactics and find a single grip they could use reliably. One participant reported this made them feel "like a claw machine" as they completed the task.

3.4.4 Maze tracing test

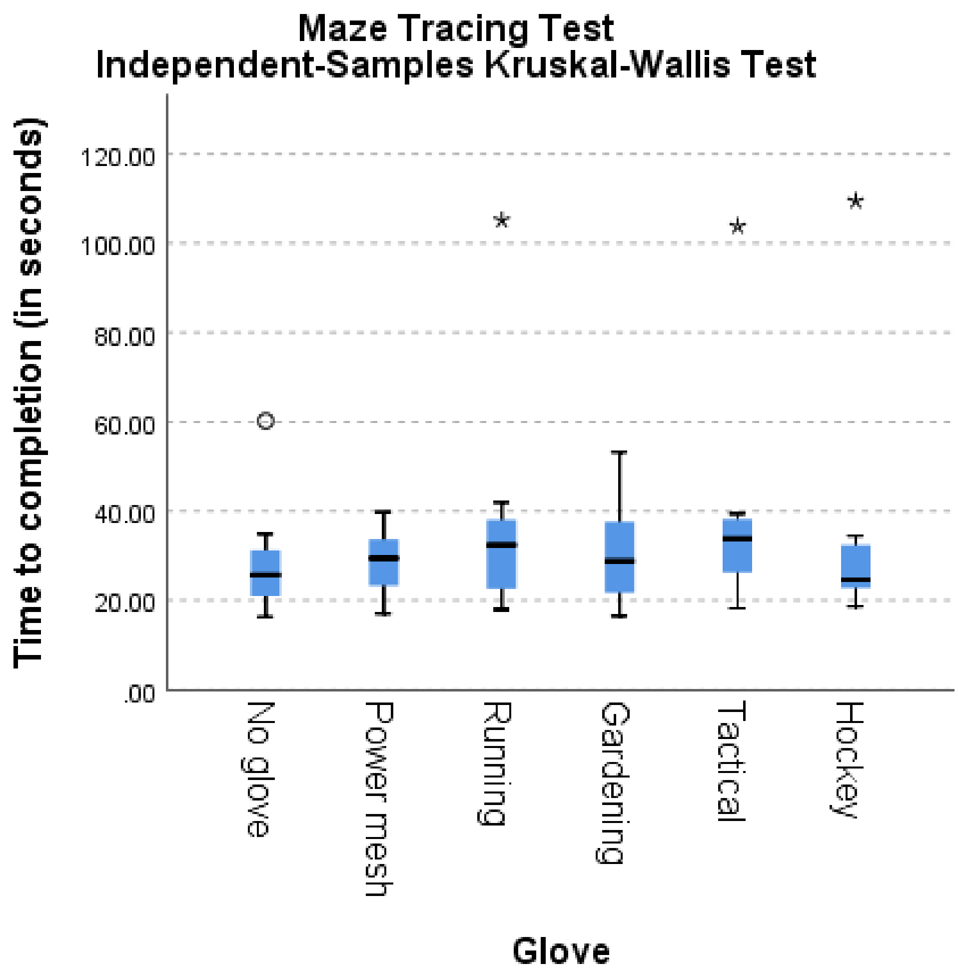


Figure 12 – Maze tracing test Kruskal-Wallis test results – there is no significant difference in the time to completion between gloves, $\chi 2(5) = 4.49$, $p = 0.481$

The maze tracing test differs from the other tests in the protocol in that most participants completed the task by primarily moving their wrists, elbows, and shoulder, rather than their fingers. They explained this as a way to maintain greater control over the pen, as they traced quickly through the maze. Despite the locked grip that most employed with the pen, they still felt the effects of the encumbrance at their fingertips – through their inability to get a secure grip on the pen, and through feeling the pen slipping out of their fingers. The results for this test did not show a significant difference in performance, as shown in Figure 12.

However, the participants did perceive a difference between the gloves in their SEQ answers – the Running gloves have a slippery smooth finish, which makes it hard to establish and maintain a grip with smooth surfaced objects. The Tactical gloves were bulky, and have stiff pads over their fingertips, which made it hard to maintain control of the pen. As seen in Figure 13, the Powermesh gloves were preferred overall, but the gardening gloves did receive praise for their superior grip, and enjoyed a small advantage in the SEQ scores.

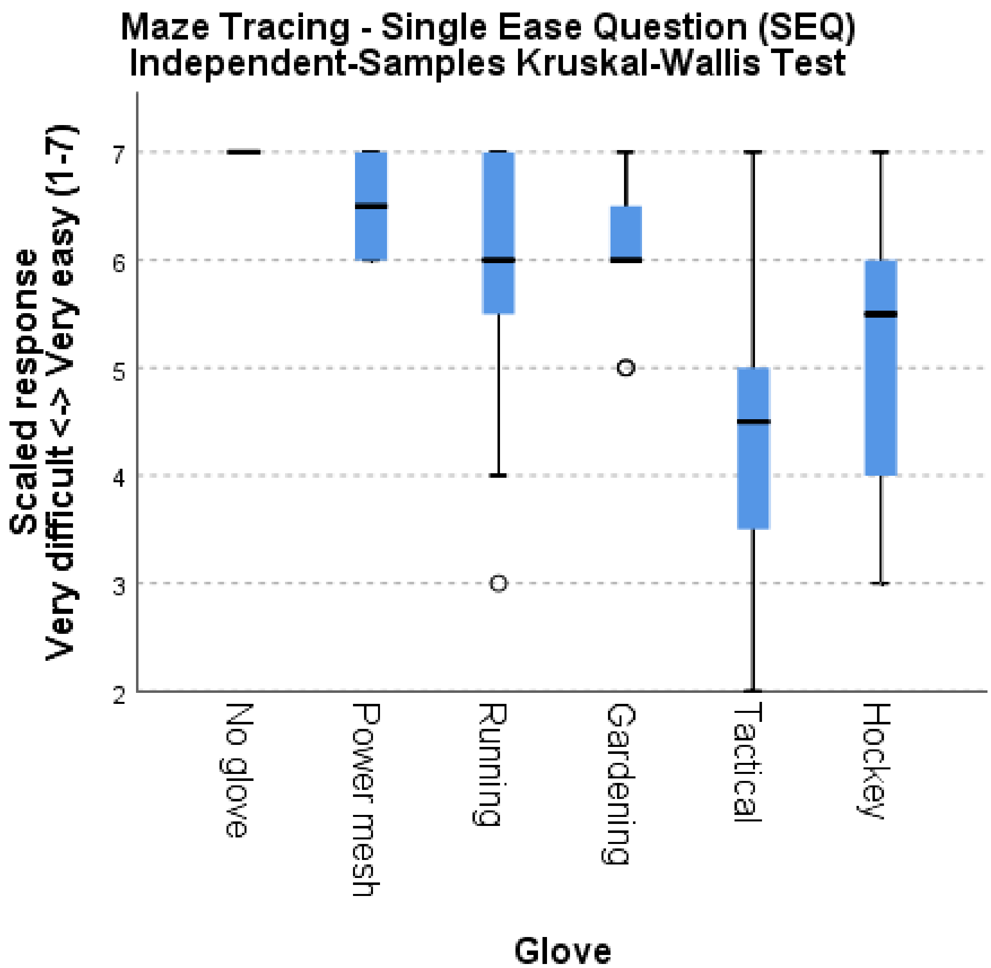


Figure 13 – Maze tracing test Single Ease Question Kruskal-Wallis test results – there is a significant difference in the perceived difficulty of the task between gloves, $\chi2(5) = 32.96$, $p = 0.000$

3.4.5 Mr. Potato Head test

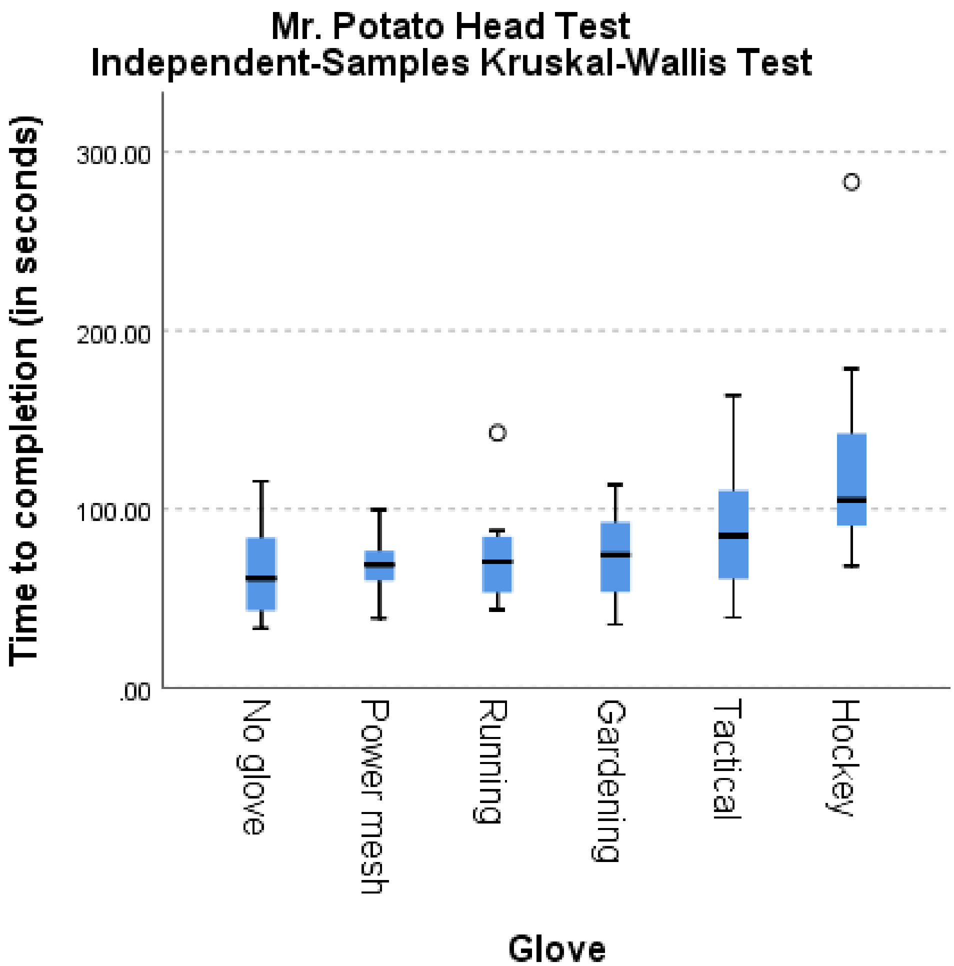


Figure 14 – Mr. Potato Head test Kruskal-Wallis test results – there is a significant difference in the time to completion between gloves, $\chi 2(5) = 17.36$, $p = 0.004$

The three lighter gloves had similar median performance times in the Mr. Potato Head test, with the Tactical and Hockey gloves showing significantly different scores in Figure 14. Participants reported that the bulk of these gloves made it harder to manipulate the parts and fit their fingers in smaller spaces.

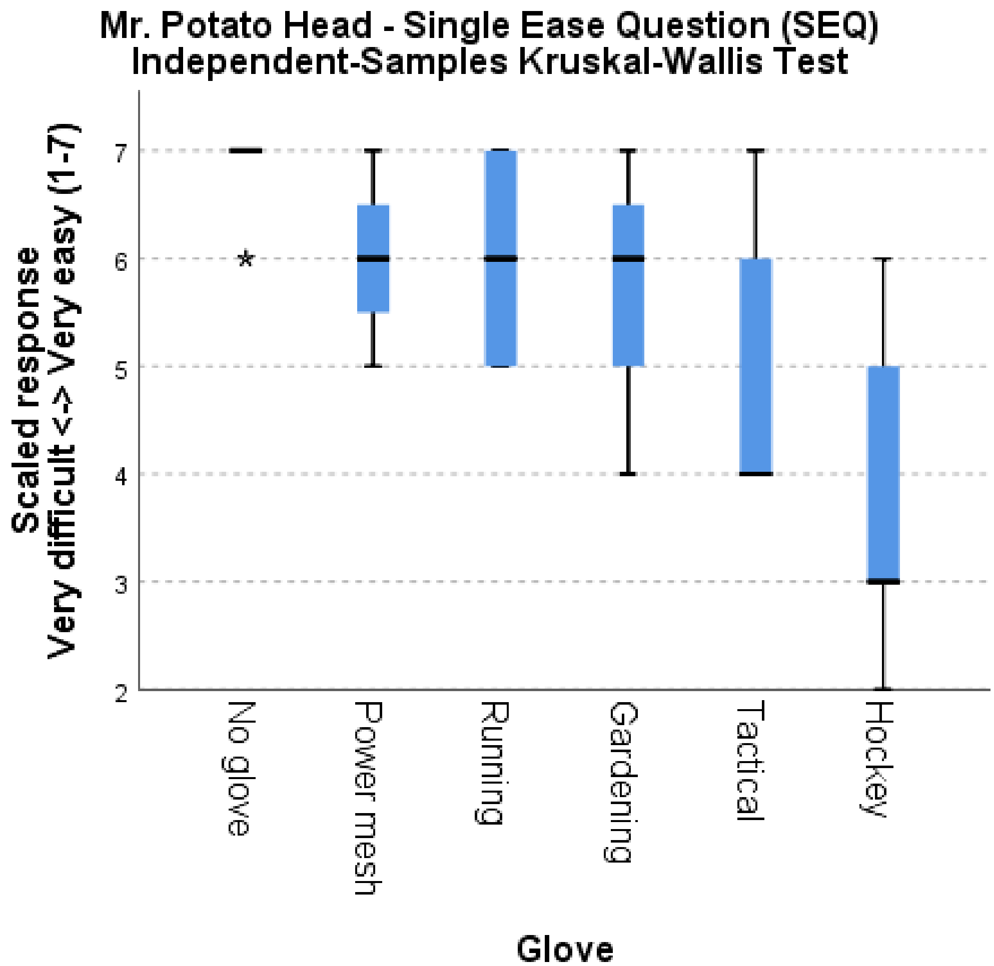


Figure 15 – Mr. Potato Head Single Ease Question Kruskal-Wallis test results – there is a significant difference in the perceived difficulty of the task between gloves, $\chi2(5) = 34.05$, p = 0.000

The SEQ scores tell a similar story in Figure 15, with the lighter gloves showing similar median scores, and the bulkier gloves perceived as making the task increasingly difficult. While the scores are similar, participants did mention material slipperiness as an effect, as they noticed the gardening glove's grippy surface did offer some benefit in maintaining control of the objects.

Participants also reported on the effects the visual occlusion of the larger gloves – wearing them made it harder to know exactly what they had picked up, and what was in their hand. This meant they needed to visually confirm many more pieces than with the smaller gloves, which slowed them down and added to their overall workload. Another effect observed was the need to use two hands to manipulate and position objects with the larger gloves – with small gloves and bare hands, most participants could move objects around just using their fingertips on one hand.

3.4.6 Geometric solids test

Participant favored the Powermesh gloves for this test. Due to the focus on sensation and perception, participants preferred the thinner gloves that provided more "sensory transparency" and were generally less concerned with additional grip, though the gardening gloves had the closest median time score – seen in Figure 16. Participants used several different strategies to find the objects while blinded – typically either matching what they felt in their hand to an image of the object they had in their head, or a checklist of features.

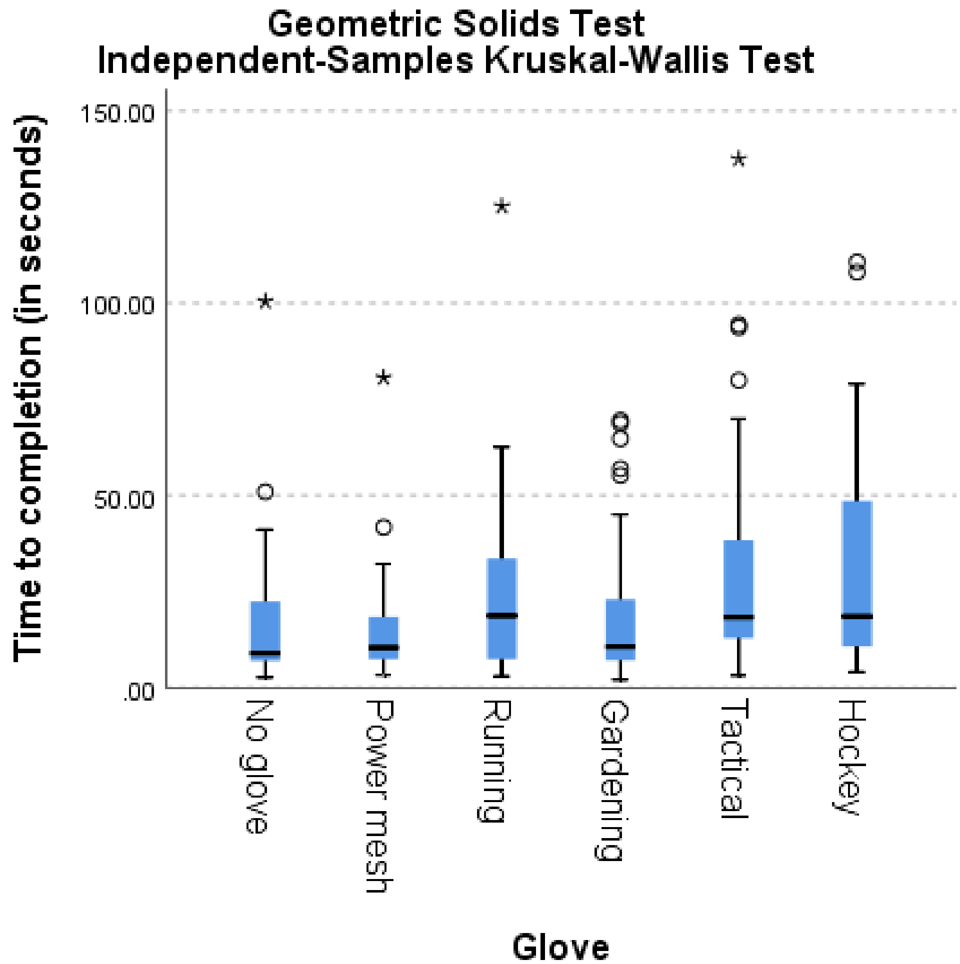


Figure 16 – Geometric solids test Kruskal-Wallis test results – there is a significant difference in the time to completion between gloves, $\chi 2(5) = 16.61$, $p = 0.005$

Not all participants found the task easier with bare hands – some even felt overwhelmed by the load of stimuli they received when concentrating during the task. Others reported a strong preference for bare hands and reported that experience felt much different than with gloves on. Many found it hard to describe the feeling, as they just somehow knew what the object was, without having to spend the time to figure it out.

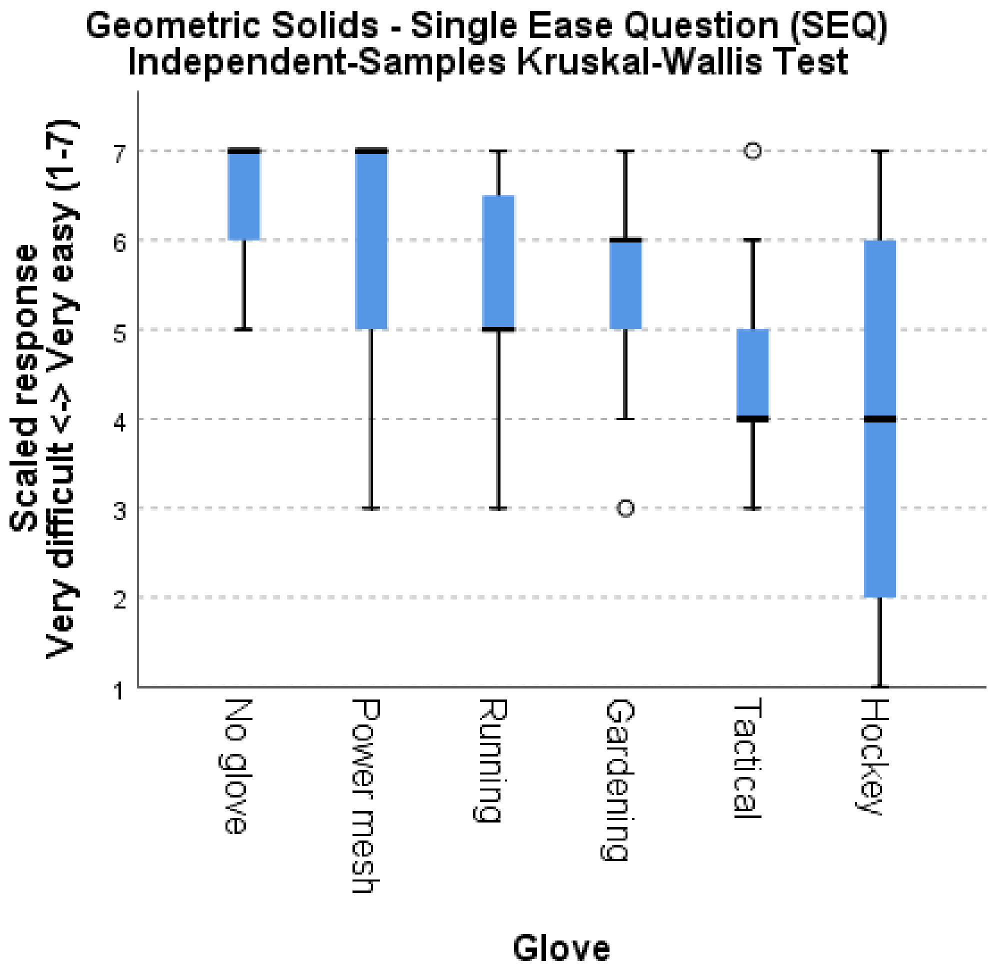


Figure 17 – Geometric solids test Single Ease Question Kruskal-Wallis test results – there is a significant difference in the perceived difficulty of the task between gloves, $\chi2(5) = 19.29$, $p = 0.002$

The SEQ scores show the wide variability in the perceived difficulty of this task across gloves – particularly the hockey gloves. The hockey gloves are lined with thin material on the palmar side, which did not greatly inhibit the participant's ability to feel and identify features. Yet the prominent padding on the dorsal side of the glove made it difficult to control and manipulate objects in the box.

3.5 Glove findings

3.5.1 Scaled response – Typing questions

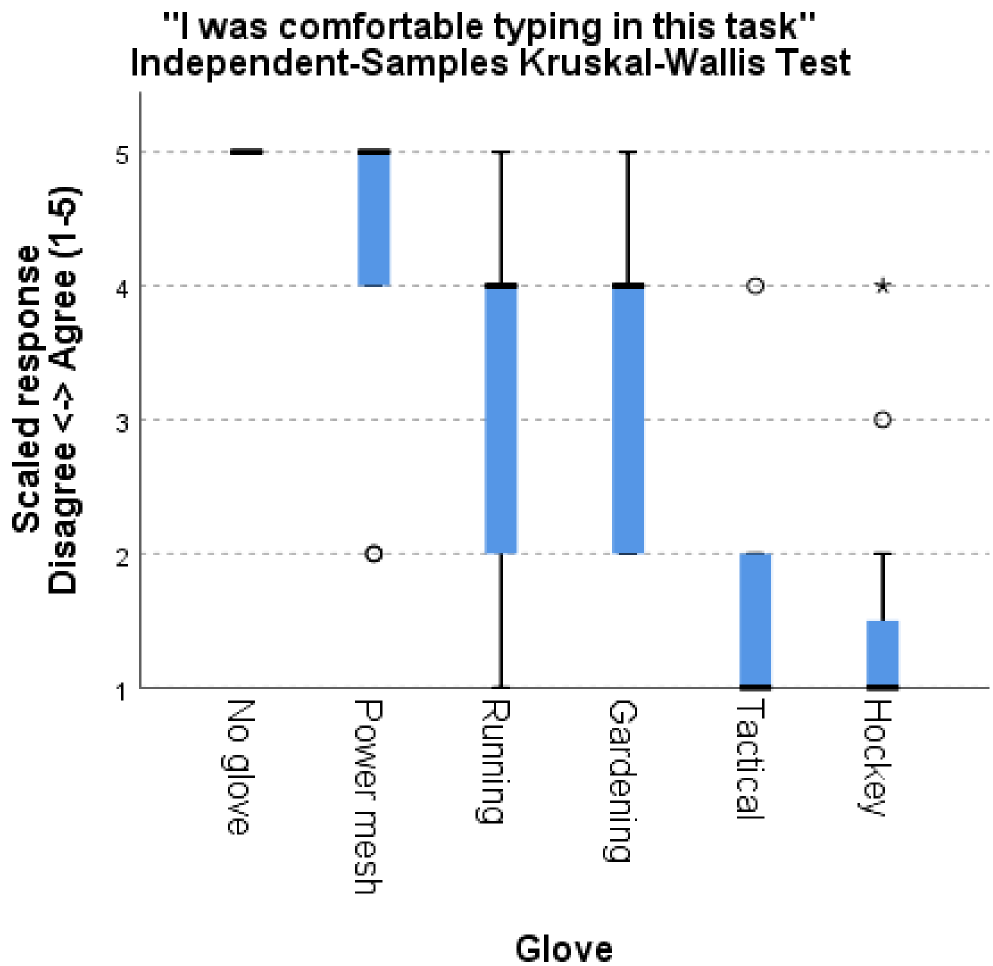


Figure 18 – Scaled response for "I was comfortable typing in this task" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi2(5) = 49.74$, $p = 0.000$

Participants were asked three questions specifically about their typing experience. The results of the first question are seen in Figure 18, which related to their comfort while typing – their responses showed a significant difference between the gloves. Participants indicated they felt comfortable typing with the first three gloves but expressed a strong indication of discomfort for the Tactical and Hockey gloves. Participants singled out the Hockey gloves for making them not able to rest their hands in a proper typing position and needing to elevate their wrists at an uncomfortable angle.

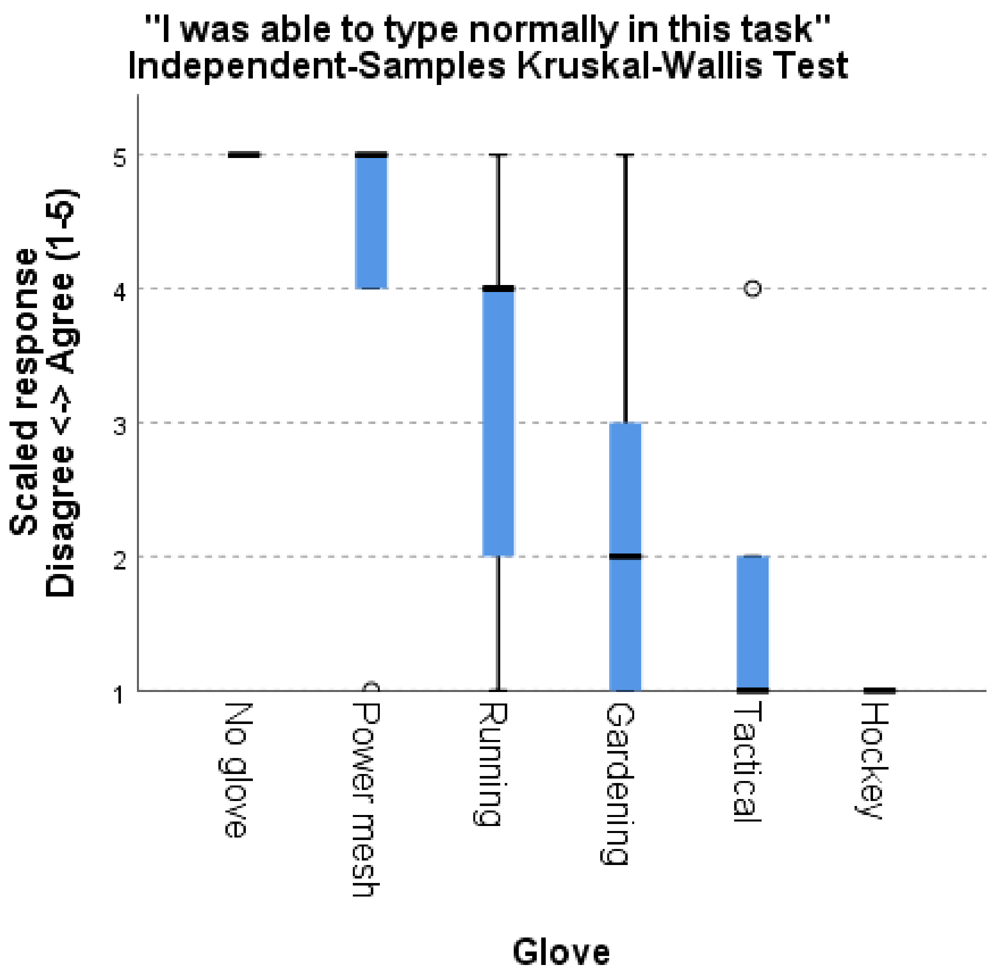


Figure 19 – Scaled response for "I was able to type normally in this task" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi2(5) = 49.61$, $p = 0.000$

Similar significant differences between gloves, seen in Figure 19, were observed in the participant's ability to type normally during the task. The Tactical and Hockey gloves again were singled out as having completely disrupted participant's normal typing abilities. The added bulk of both gloves made it hard for participants to feel the keys, differentiate between adjacent keys, and feel whether they had successfully pressed a key. Additionally, the fingers on the Hockey gloves were larger than the keys on the keyboard, which participants reported made it very difficult to hit a single key.

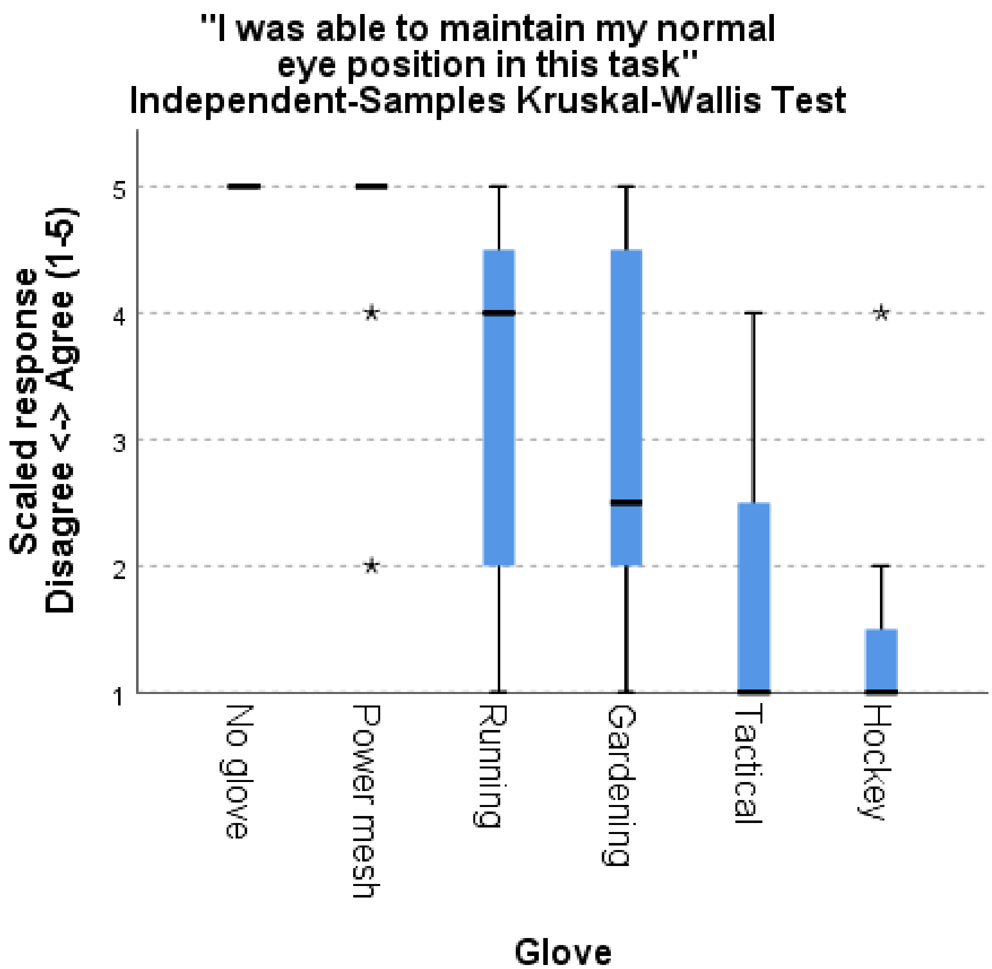


Figure 20 – Scaled response for "I was able to maintain my normal eye position in this task" Kruskal-Wallis test results – in the typing task there is a significant difference in agreement with the statement between gloves, $\chi2(5) = 46.55$, $p = 0.000$

The final question about the participant's typing experience related to their normal eye position. This result again showed a significant different between the gloves, as the Powermesh gloves showed no difference to the participants gaze than the "no glove" baseline condition. However, the Hockey glove again totally altered participant's typing capabilities, which changed the position of their gaze. As participants could not feel the keys, they needed to visually confirm whether a key had been struck by looking at the keyboard. However, the Hockey glove occluded their ability to see the keys, which meant they need to swing their gaze back to the screen to confirm they had successfully typed the correct character. This resulted in an oscillating gaze for many participants, who needed to keep their gaze in constant motion to confirm they were correctly completing the word. Participants reported that this increased the mental effort required to complete the task.

3.5.2 Scaled response – Comfort and fit

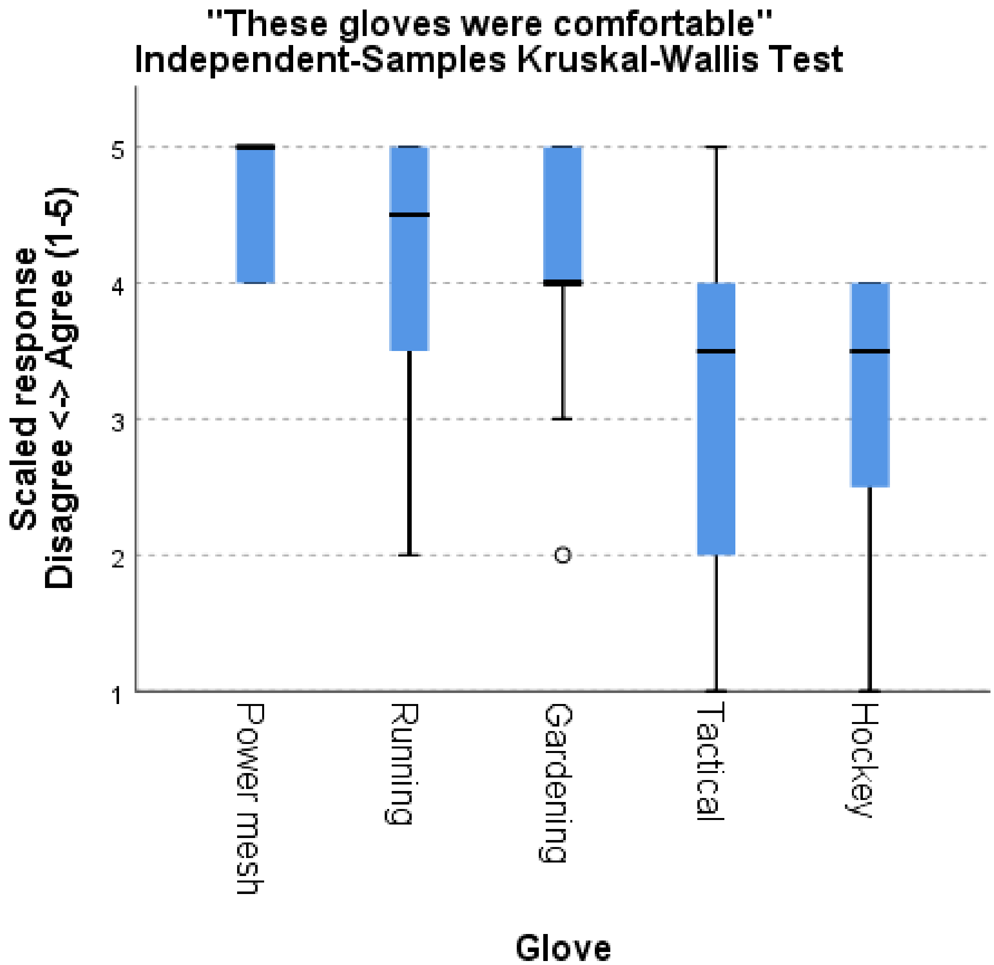


Figure 21 – Scaled response for "These gloves were comfortable" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi2(4) = 17.365$, $p = 0.002$

Perceived glove comfort was shown to be significantly different between the gloves, as shown in Figure 21. Participants singled out the Gardening, Tactical, and Hockey gloves for heating their hands during the tasks and leading them to feeling sweaty. The Powermesh glove, by contrast, were mentioned as keeping their hands cool.

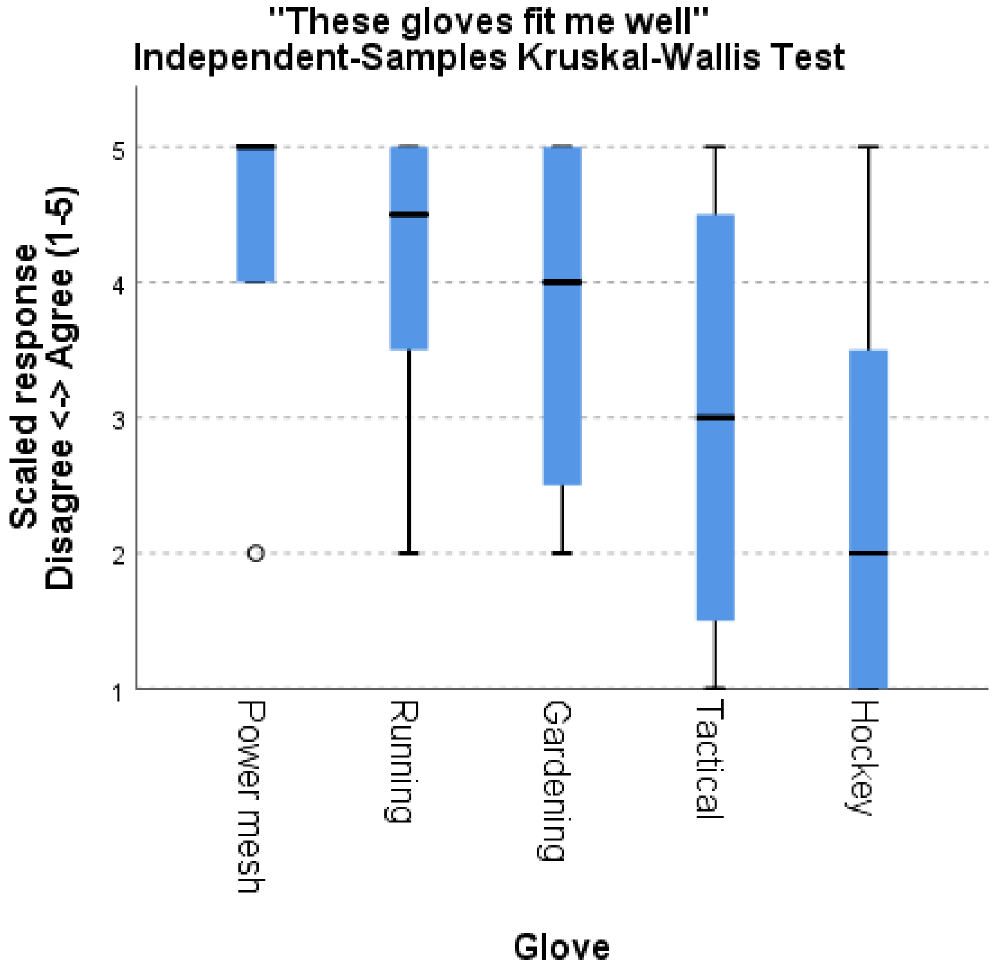


Figure 22 – Scaled response for "These gloves fit me well" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi2(4) = 16.47$, $p = 0.002$

Perceived fit showed a similar significant difference between gloves – seen in Figure 22. Fit has a notable impact on task performance, as tight gloves were reported to restrict participant's movements, and loose gloves bunched and got in the way – especially at the fingertips. The range of the response for the Tactical gloves is likely due to the wide ranges of sizes available, which likely helped participant's tailor their fit. By contrast, the Hockey glove was available in a single size, and was a poor fit for most participants.

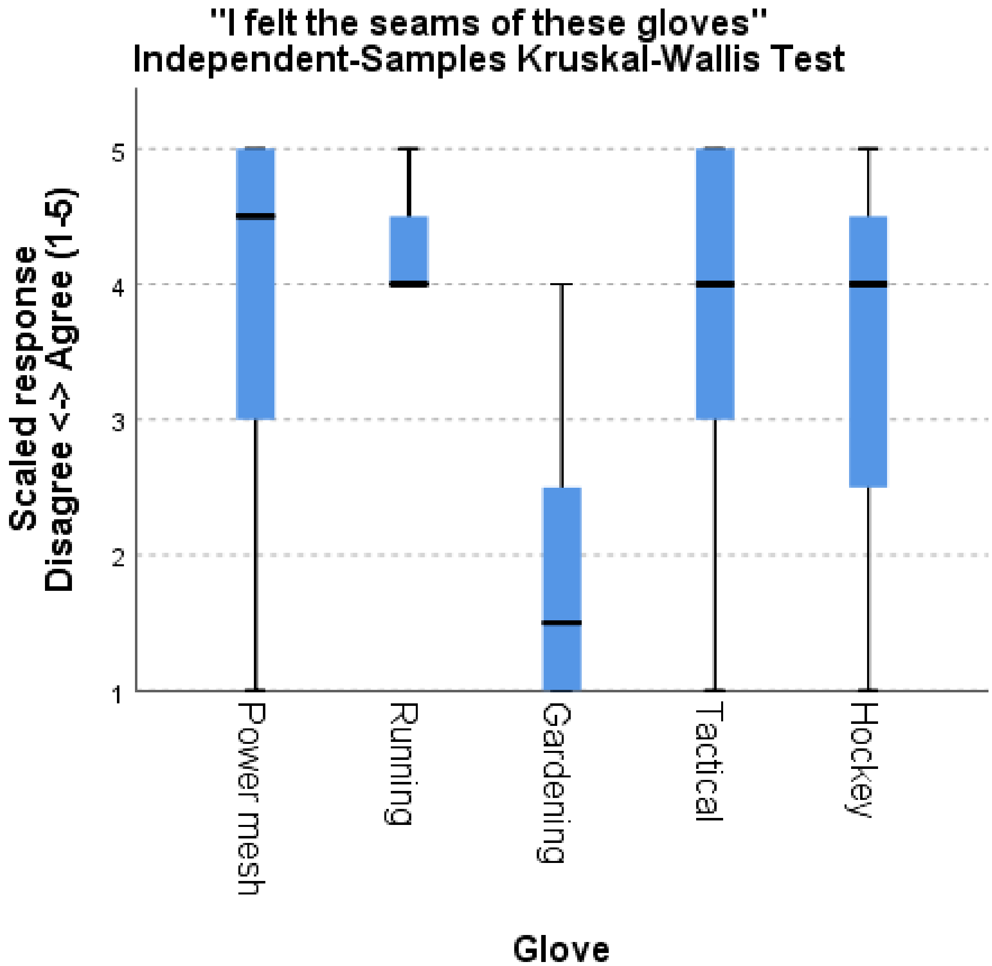


Figure 23 – Scaled response for "I felt the seams of these gloves" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi2(4) = 18.70$, p = 0.001

Seams were singled out in early pilot questions, and the results showed a significant difference between gloves in participant's ability to feel them. The results show the Gardening glove as the only one of the set that did not have prominent seams – due to continuous knit construction. Participants reported feeling the gloves seams as being particularly noticeable in the monofilament test, and the box and blocks test – both of which had them focusing on small details at their fingertips. However, there were reports of the seams getting in the way in both the box and blocks task, and the Mr. Potato Head task – especially with the long floppy fingers of the Running glove, and the oversize fingers of the Tactical glove. Small details of glove construction could have big effects on participant's perception of small details and help them to feel that they had full control of their fingers.

3.5.3 Scaled response – Slip and grip

Slip and grip were frequent topics of discussion, and the results for this question show that there were significant differences in the perceived slipperiness of the glove surfaces. The Powermesh and Running gloves were consistently described as slippery, given the smooth finish of their surfaces. By contrast, as shown in Figure 24, the Gardening glove was seen as the least slippery, due to its Nitrile dip finish.

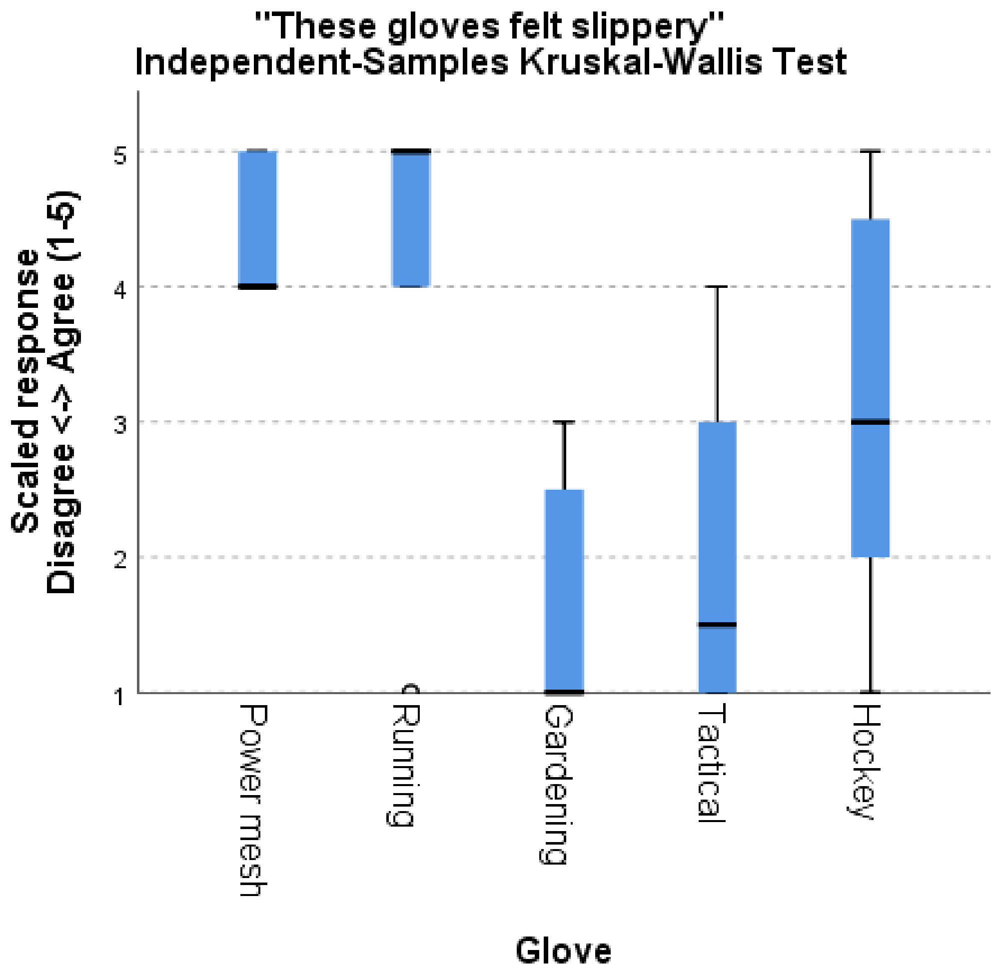


Figure 24 – Scaled response for "These gloves felt slippery" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi 2(4) = 30.38$, $p = 0.000$

The question of glove grip showed the inverse of the slip question, with significant differences between the gloves shown in Figure 25. Participant's observations again broke the gloves down into three groups. The gardening gloves were singled out for their superior grip, the Power mesh and running gloves were reported to feel slippery across the tasks, and the Tactical and Hockey gloves were in the middle – though participants reported them to be both slippery and grippy, due to other encumbering effects with their large fingers.

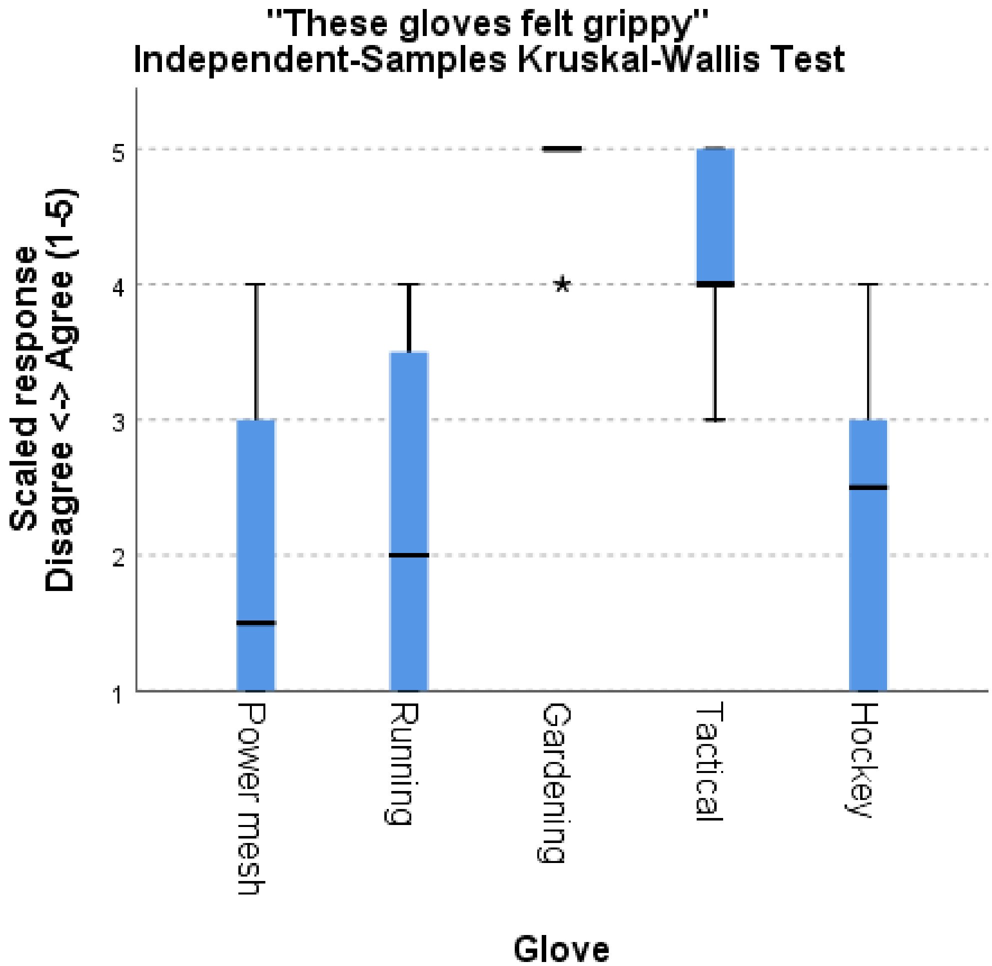


Figure 25 – Scaled response for "These gloves felt grippy" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi 2(4) = 39.82$, $p = 0.000$

3.5.4 Scaled response – Hand mobility

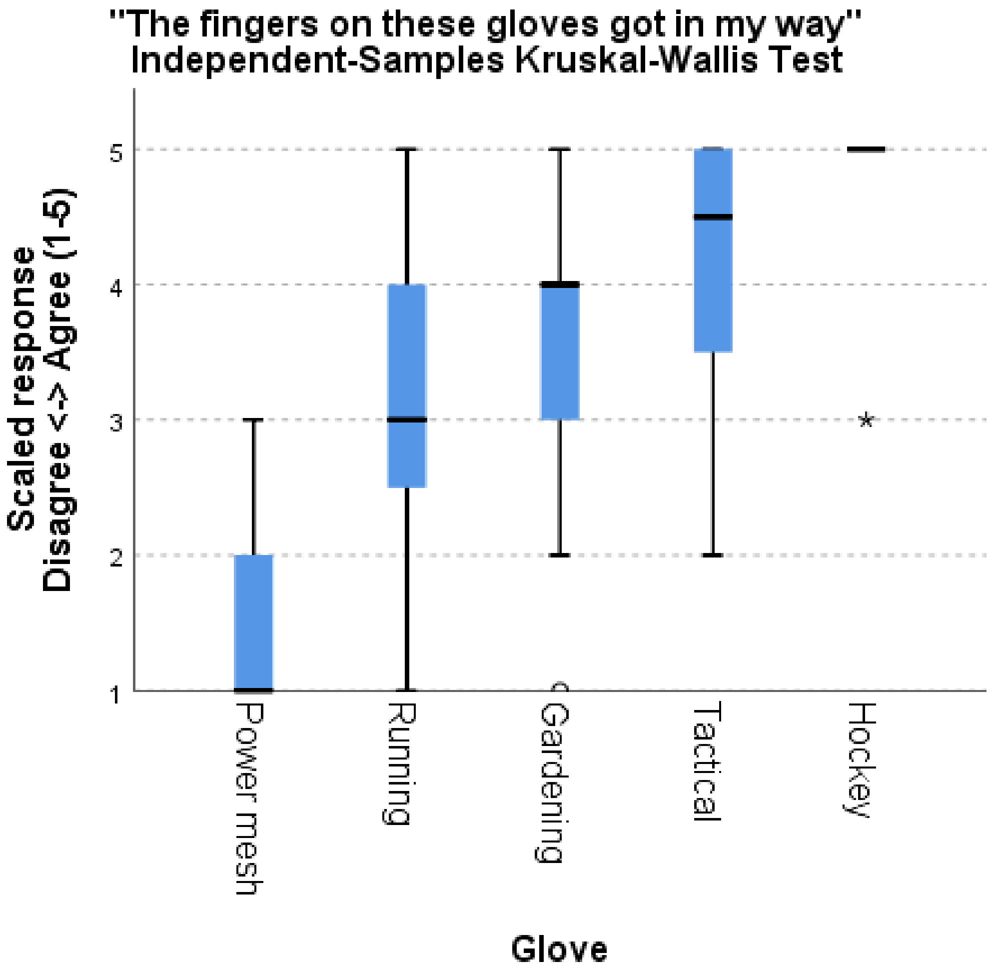


Figure 26 – Scaled response for "The fingers on these gloves got in my way" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi2(4) = 33.74$, $p = 0.000$

The final set of questions focused on the glove's impact on the participant's hand mobility. The results, seen in Figure 26, show a significant difference in how participant's perceived the gloves getting in their way during the tasks. The Powermesh gloves scored very well, as there were few ways they were thought to interrupt the participant's interaction with the task objects. The Hockey gloves bulk and poor fit led many participants to report that they had completely interrupted their tactile and sensing capabilities. The long fingers were particularly detrimental to many participants, who felt like their fingers were no longer useful for objects manipulation.

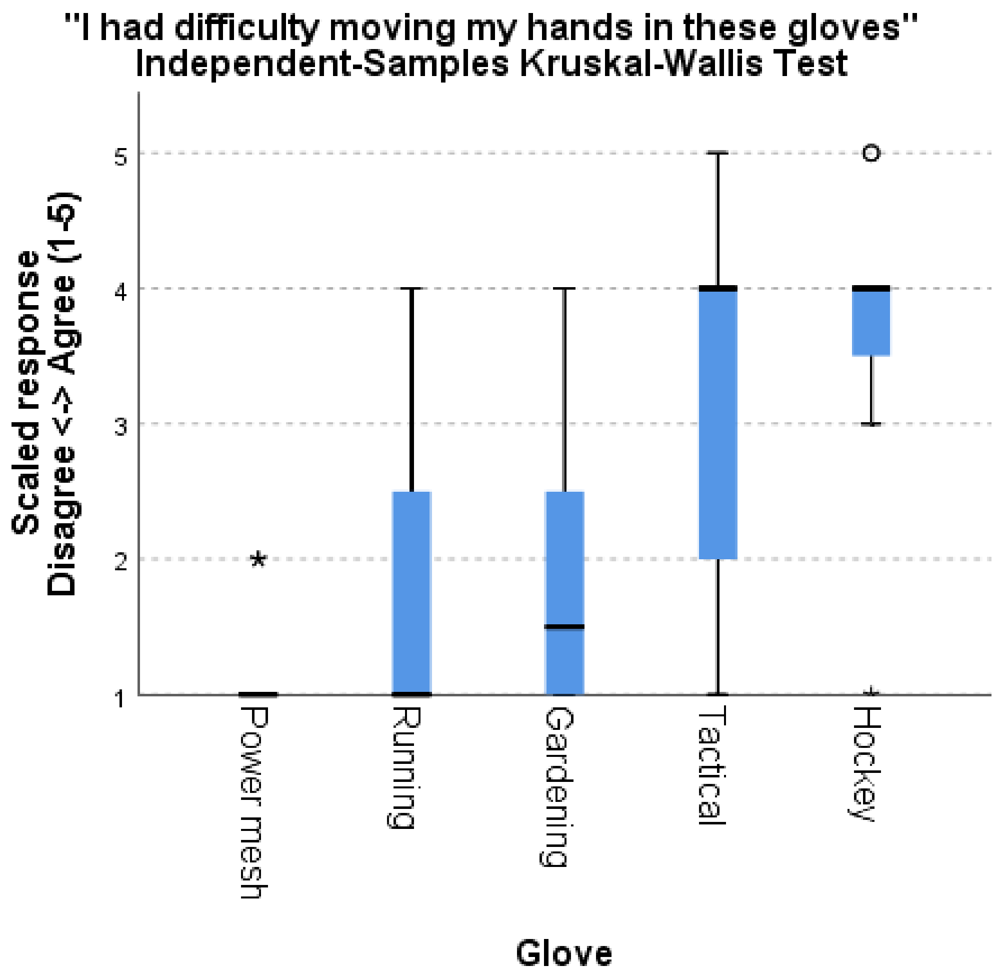


Figure 27 – Scaled response for "I had difficulty moving my hands in these gloves" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi2(4) = 30.05$, $p = 0.000$

The final question related to perceived impact on hand movements, and there was a significant difference shown between the gloves. The Tactical gloves performed poorly in this assessment, with participants reporting that their stiff materials and constricting fit made it difficult to move their fingers.

3.6 Discussion

3.6.1 Glove preference

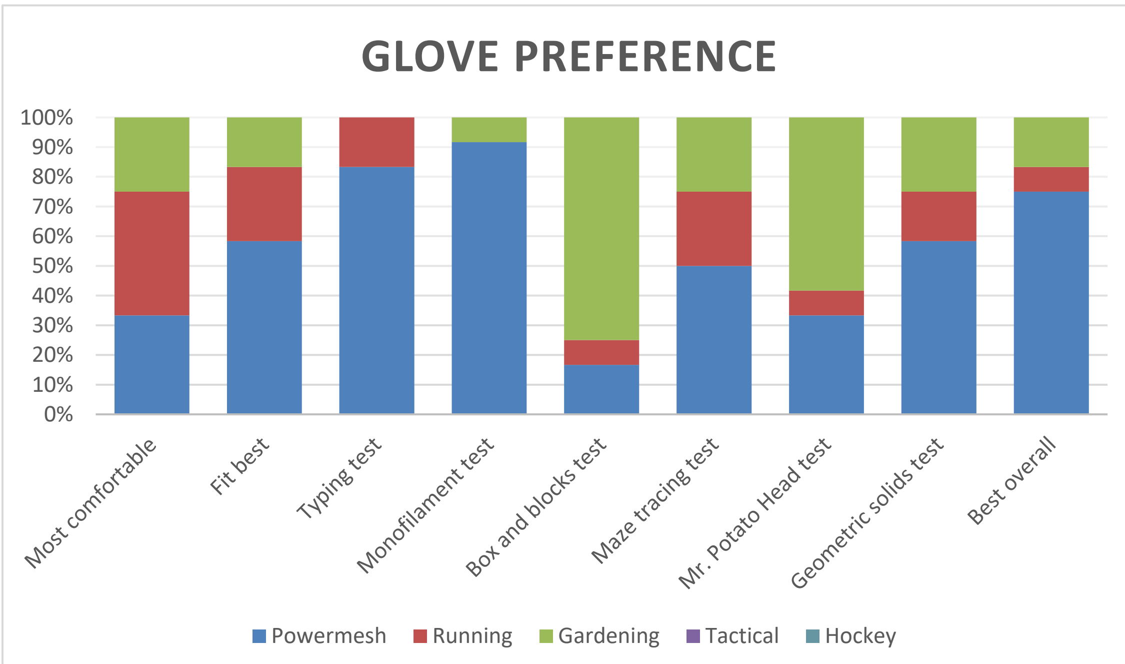


Figure 28 – Aggregated answers for post-session interview questions about preferred glove choice for comfort, fit, and task completion.

Participants showed a strong preference for the Power mesh glove for both fit and comfort, which led to a dominant performance in six of the nine preference questions, including the best overall. As shown in Figure 28, the Powermesh gloves were perceived as lightweight and comfortable, though the Running gloves were preferred by a slight margin, for their soft finish. Beyond comfort, the Powermesh gloves were preferred due to the sensory transparency of the fabric – participants were able to get more information from feeling the objects.

Table 3 – Aggregated values for post-session interview questions, with totals. Values shown indicate the number of participants who picked each glove as their preferred option for each question about comfort, fit, and task completion.

Preference	Powermesh	Running	Gardening	Tactical	Hockey
Most comfortable	4	5	3	0	0
Fit best	7	3	2	0	0
Typing test	10	2	0	0	0
Monofilament test	11	0	1	0	0
Box and blocks test	2	1	9	0	0
Maze tracing test	6	3	3	0	0
Mr. Potato Head test	4	1	7	0	0
Geometric solids test	7	2	3	0	0
Best overall	9	1	2	0	0
Total	60	18	30	0	0

The only contrast to this result was the performance of the Gardening gloves with the Box and Blocks and Mr. Potato Head tests, due to their superior grip. This continues the observations of preference and performance for the Gardening gloves when grip and object control was the primary concern of the participants, and the Powermesh when sensory transparency and comfort were the primary concerns. These results suggest that participants would prefer a set of gloves that performed similarly to the Powermesh gloves, with the added grip capabilities of the Gardening gloves.

3.6.2 Cognitive Load

Cognitive load was not tracked precisely through this study, but the discussions following each task made it clear that participants were experiencing an increased mental workload as the encumbrance effects of the gloves increased. Much of this can be attributed to task adaptation, as participants reported needing to think about how to complete the task in new ways, when the glove encumbrance made their first choice more difficult.

3.6.3 Visual Confirmation

Another area where participants needed to work harder was with an increased demand on visual confirmation. This effect was noticed when participants could not tell what was in their hand through touch alone and could not see what was in their hand due to the occlusion of the glove. This was more common with the larger gloves and meant that participants spent more time rotating their hands and visually confirming what was in them, than with the smaller and lighter gloves.

3.6.4 Conclusion

It is clear that wearing gloves can have a significant impact on a person's ability to complete tasks that involve object interaction. Any gloves built for Mixed Reality contexts that require interaction with both virtual and physical objects will require careful design and testing to ensure that the user has equal unencumbered access to both. This study showed some of the potential effects that will need to be addressed, and the results suggest some initial opportunities for design research to investigate.

Adding grippy material to the fingertips of gloves seems to provide a big benefit to the user experience and should likely be considered for any future glove product. Keeping seams and excess bulk away from the fingertips and distal phalange of the finger also increases the chance that the user makes good contact with the surface of the object, dig their fingers into small spaces, and maintain visual contact with the object in their hand – all of which should improve their task performance. This study also highlighted the necessity of considering users with long fingernails in the design of interactive gloves, and the unique encumbrance challenges they will face.

Encumbrance effects will always slow users down and make them work harder to complete the same tasks. Encumbrances may be unavoidable, but they can be managed and designed for. One of the best ways of understanding the encumbrance effects that participant experienced in this study came from their description of completing the typing test with no gloves on, after the same task wearing gloves. One participant described their experience like this:

"It feels like swinging a bat with one of those practice weights on it, and then getting to bat normally"

Another participant was more succinct:

"It feels like going home again."

CHAPTER 4. DESIGN OF THE COST OF HAPTICS PROTOCOL

The Glove Encumbrance Study showed the potential of testing glove encumbrance effects through task-performance metrics, and the benefit of running a mixed-methods study. It became clear that the quantitative measures were capable of showing the effect size of the encumbrance, which allowed the qualitative measures to suggest the possible reason for the effects. While the study results generated useful data and a set of design recommendations, it was clear that there were many areas that could be improved and enhanced. This proposal outlines a series of studies, that will be undertaken to gain further detail and insight into the nature of glove encumbrance.

4.1 Study Updates

The new study program, named *The Cost of Haptics*, was designed to address a few key opportunities that arose from the review of the previous study.

4.1.1 Test Validity

The tasks selected for the preceding study were inspired by common tasks used by Occupational Therapists in the clinic and seemed to perform well in providing a functional range of tests for hand interaction. However, as the focus of the new study shifts to more generalized examination of glove encumbrance, there is an opportunity to select tasks that have been repeatedly reviewed and been shown to be valid in the broader literature. This helps address concerns that the tasks selected in the previous study are capable of showing these encumbrance effects outside of the specific testing conditions of the study.

To address this, Dianat et al's literature review of glove evaluation methods for hand function was reviewed, and a short list was developed from the most common methods used by the researchers covered in the review [19]. Dianat identified at least ten different classes of hand function that could be evaluated, each with an established set of methods used to evaluate them. Of these ten, five were directly relevant to the tasks used in the previous study: manual dexterity, tactile sensitivity, fatigue, discomfort, and finger movements. The tasks associated with each class were compiled and reviewed for validity. A short list was then created of common tasks, and evaluation kits were purchased for each of them. Following a round of internal testing, the final four new tasks were then selected – they will be discussed in greater depth later in this paper.

The criteria for selection included the longevity of the tests in the research community, and the continued reviews of the function and validity. Another factor in the selection was the availability of benchmark scores showing the performance distribution of the tasks across various populations. Having access to this data – which wasn't available for most of the previously selected tasks – will help ensure that the results from the new studies can be accurately placed against data from the broader population, which will greatly benefits any questions of validity. This is also important for the broader focus of the new study on benchmarking the results.

4.1.2 Independent variables

The original intent of the prior study was to determine if task-performance tests could show differences in encumbrance effects between gloves. This meant that purchasing commercially available gloves was a strategy suitable to answering the research question, so long as they were arranged along an approximate axis of expected encumbrance. While this was sufficient to prove out the basic performance of the evaluation system, the gloves had many disparate design features that contributed to the overall effect for the user. There is value in measuring this amalgamated design – to test gloves against each other or standard benchmarks – but getting detailed answers about the actual effect size of specific design features requires a more precise approach.

This leads to the proposal of creating gloves for the purposes of testing these features – each glove set presenting a specific independent variable that can be manipulated and tested for effect. These independent variables would be drawn from essential glove design attributes – material thickness, texture, stretch, hardness, or combination thereof – and represent the most basic design features being evaluated. The performance outcomes of these tests would begin to influence the design requirements of any glove project, allowing some certainty in advance of the effects on encumbrance of any material or pattern choice being considered.

Additionally, the creation of gloves to test independent variables allows for the most detailed evaluation of high-impact design features. In circumstances where there is a high-risk choice between two competing design alternatives, constructing gloves that isolate the features for testing will provide additional detail that can better inform the decision. This allows for testing early enough in the design process to save time and costly revisions, compared against development cycles where testing only occurs after the design is locked.

4.1.3 Cognitive load

The previous study showed numerous examples where participants struggled to adapt to encumbrance effects imposed by the gloves and took on an increased cognitive burden in addition to a physical one. This adaptation asks participants to take on an additional load to think about tasks they would not normally need to think about and contributed to the feeling they were slowed down by the gloves. It is clear that measuring cognitive load is a necessary and revealing part of evaluating glove encumbrance.

Cognitive load is assessed in several ways, many still under development. There are physiological tests, such as measuring pupillary response and EEG signals, which may be useful to consider for future studies, but they provide their own encumbrances with the measurement apparatus for the user [71]. A better fit for this study is likely one of the subjective assessment tools, which ask users to indicate their relative feeling of mental load on a scale. This is a similar approach to the existing Single Ease Question, which is already being administered post-task. There are two main options which were considered for this measure – the Paas cognitive load scale, and the NASA-TLX Mental Demand [31,65].

4.2 Task Performance Measures

The new study requires a set of task-performance measures that meet the discussed criteria for validity and cover a range of hand functions. Following the identification of the short-list of tasks from Dianat's review, a new set of tasks were evaluated in three categories.

Grouped in a tool-use category, the Bennett hand tool dexterity test and the standard typing test were both selected, and both were added to the final list [8,82]. Tool use successfully demonstrates a range of hand function, and these tests are well established to provide valid measures. The second category deals with object manipulation and finger dexterity – the short list there included the O'Connor Finger Dexterity Test, the Functional Dexterity Test, the Minnesota Manual Dexterity Test, and the Box and blocks test [1,9,18,52]. Of these, the O'Connor and Minnesota tests were selected for this proposal. The Functional Dexterity test is effectively a smaller version of the Minnesota test, and the expanded test seems better suited to the effect sizes this study expects to be measuring. The Minnesota test was also selected in favor of the box and blocks test, as it allows for more detailed object interactions (flipping, one-hand manipulation), rather than just the speed of grasp and transport – this again seems a better match for the focus of the next study.

The final grouping were tasks that dealt with sensing at the fingertips. These included the previously used Semmes-Weinstein Monofilament test, the Cochet-Bonnet Aesthesiometer, the Two-point discrimination test, and the STI2 Shape-Texture-Identification Test [6,7,29,46]. The Cochet-Bonnet test is a similar filament test to the Semmes-Weinstein test, but varies the length of the filament, rather than the diameter, to vary the applied force. It is primarily used for ophthalmological treatment and was determined to offer no increased benefit, while being harder to apply consistently when used on fingertips. The Two-point discrimination test is a valid measure for subjects with nerve damage or some similar loss of cutaneous sensation, but it proved to be not sensitive enough for the encumbrance effects imposed by the gloves in this study. Ad-hoc testing showed a floor effect in the data, as the gap sizes needed to show the effect were smaller than the instrument allowed for. This meant that the two tests selected from this group for the next study were the Semmes-Weinstein Monofilament test and the STI2 Shape-Texture-Identification Test.

The six tasks selected therefore cover a broad range of hand activity and offer an equally wide opportunity to measure various encumbrance effects.

4.2.1 Typing test

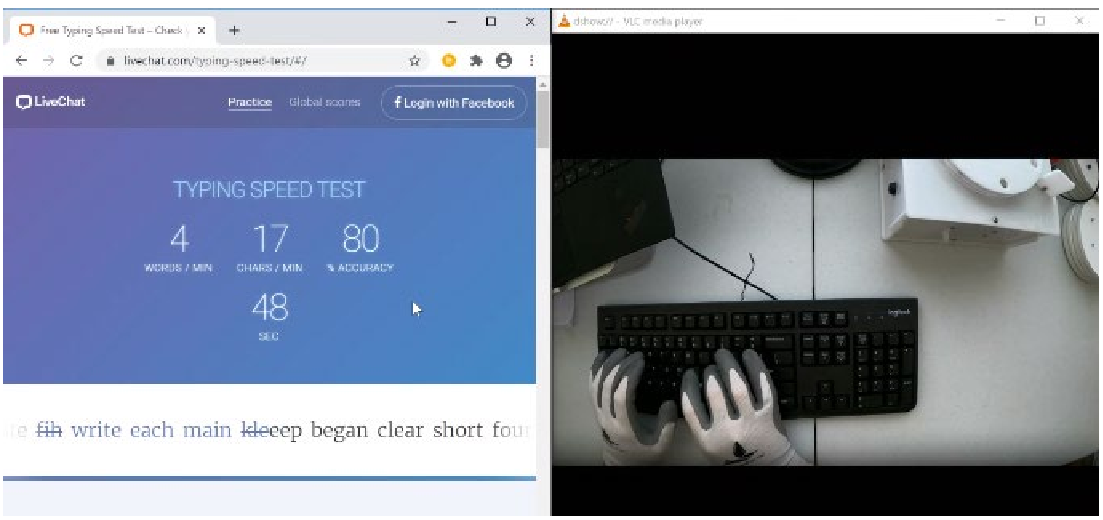


Figure 29 – A view of the typing test in action, during a remote study session. On the left is a live view of the online typing test interface, and on the right is a camera feed of the participant's hands.

Typing again offers many benefits for this evaluation, as it is a skill that takes advantage of dextrous tool manipulation and does not need to be taught to most participants. This makes it easy to measure the change in performance across a “words per minute” score, and an error rate, and infer that the gloves are the main cause for any change. The test will again employ a common online tool that uses random words pulled from the top 1000 most common words in English and asks participants to complete as many as possible correctly in 60 seconds. The task scale tests perception of fine object detail at the fingertips, and gross finger movement [63]. The apparatus, seen in Figure 29, features a Logitech k120 keyboard, and an online typing test [47].

4.2.2 Bennett hand tool dexterity test

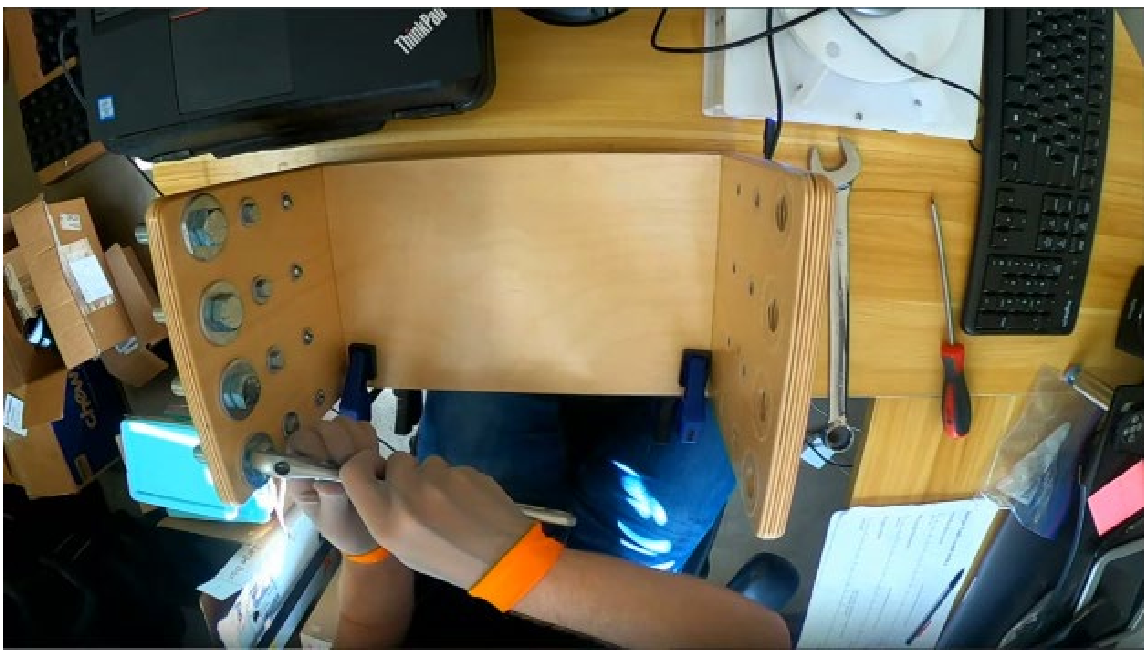


Figure 30 – The Bennett hand tool dexterity test apparatus

The Bennett hand tool test measures tool proficiency and two-handed object manipulation, ranging in scale from small mechanical nuts to large whole-arm movements with a crescent wrench and screwdriver [8]. The participant is tasked with removing 12 sets of nuts and bolts and transferring them from one side of the apparatus to the other, as shown in Figure 28. The total task time is measured and can be compared to a reference chart showing performance across various populations. This task has been in continuous use since the 1950s and is well-documented.

4.2.3 O'Conner Finger Dexterity Test

Figure 31 – O'Conner Finger Dexterity Test apparatus

The O'Conner Finger Dexterity test, seen in action in Figure 29, is suited for fine dexterity object manipulation, asking participants to sort 300 pins from a container into 100 slots, 3 pins in each slot. This task requires fine control and relies on detailed sensing to manipulate 3 small objects simultaneously. Performance in the task is measured with total time to completion, though there are options available to weight the second half of the task to account for fatigue. This task also has well documented performance data, which will allow for easy comparison of results against the general population.

4.2.4 Minnesota Manual Dexterity Test

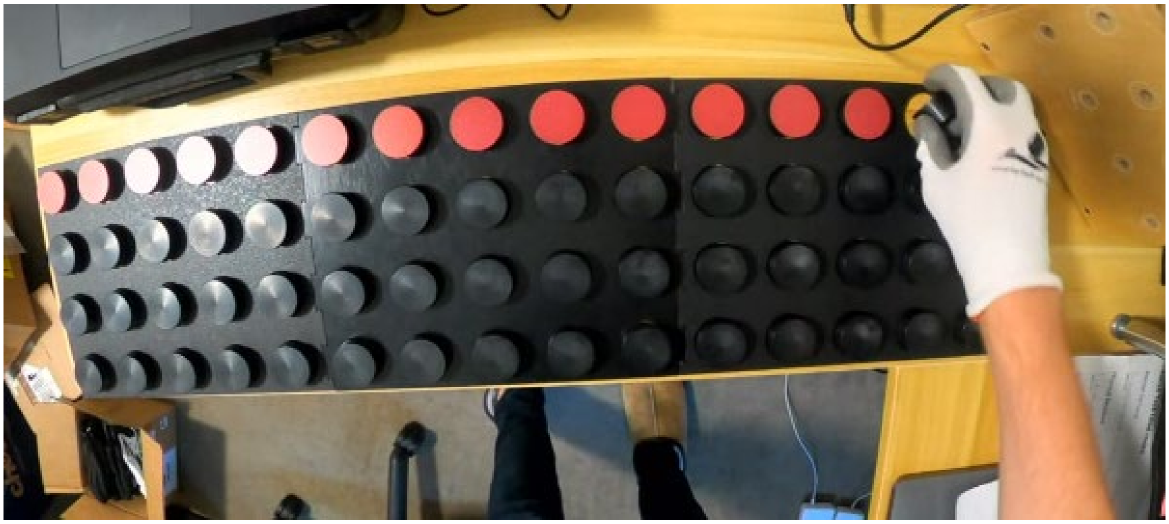


Figure 32 – The Minnesota Manual Dexterity Test apparatus

The Minnesota Manual Dexterity Test allows for testing of rapid object manipulation and placement. As seen in Figure 30, the participant is asked to move or flip 60 two-color disks in a grid, using either one hand or two, with a total time to completion indicating their performance. In this study, the one-handed flip variation allows for examination of in-hand object manipulation as part of a pick and place task, which asks for relatively complex independent finger movements, and fine control of surfaces. The 60 repetitions should make performance differences readily apparent between different conditions and provide a well validated task for measuring medium-scale objects.

4.2.5 Semmes-Weinstein Monofilament test

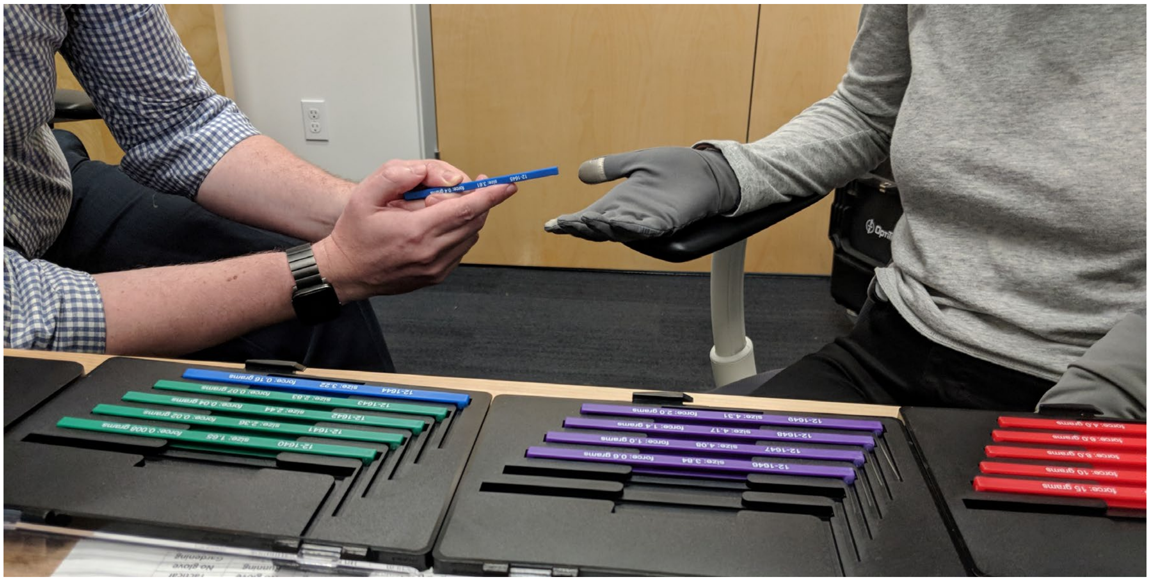


Figure 33 – The Semmes-Weinstein Monofilament test apparatus

As in previous studies, the Semmes-Weinstein Monofilament test provides the easiest and most valid way of measuring the detection of the smallest force that can be perceived against the fingertips, and any changes in perception due to the various glove conditions. While there are more precise ways of measuring these effects, the monofilament test provides an affordable and effective way of measuring these differences and performed well in the previous studies. The task is blinded for the participant, where a filament is placed on a finger, and bent to a 90-degree angle – as shown in Figure 31. The participant then indicates if they feel anything, and which finger they believed detected the force – in this study the thumb, index finger, and pinky will be evaluated. The test continues until an obvious detection threshold is reached, with three correct confirmations in a row.

4.2.6 Tactile Discrimination Test

Figure 34 – The Tactile Discrimination test apparatus

The Tactile Discrimination test is designed to test the glove's relative impairment of the participant's haptic gnosis – their ability to gain information about the world through their sense of touch. This test features 6 shapes in three pairs – circle and oval, square, and hollow square, star and hexagon – each of which test difference shape identification characteristics. These shapes are arranged on plates, and are available in nine different scales, ranging from 2mm to 10mm in diameter for the individual shapes. Participants are presented the plates while blinded and asked to identify the shape they are feeling with respect to a visual reference. The smallest size the participant is capable of correctly identifying for each shape is recorded as the score for the task. For more information on the development of the Tactile Discrimination test, please refer to Chapter 5.3.

4.3 Subjective Measures

Subjective measures proved to be vital in previous mixed-methods studies, as they offered possible explanations for the effects shown by the quantitative measures.

4.3.1 Perceptual Scaled Measures

These tools are useful for evaluating user's perceptions of difficulty and load in the tasks, and provide scaled response data that is straightforward to analyze. NASA-TLX is the best known instrument in this set, and is a useful way to sample a broad range of perceived loads across the task [30]. The scales measured by NASA-TLX are as follows:

1. Mental Demand – How mentally demanding was the task? [0 Very Low – 21 Very high]
2. Physical Demand – How physically demanding was the task? [0 Very Low – 21 Very high]
3. Temporal Demand – How hurried or rushed was the pace of this task? [0 Very Low – 21 Very high]
4. Performance – How successful were you in accomplishing what you were asked to do? [0 Perfect – 21 Failure]
5. Effort – How hard did you have to work to accomplish your level of performance? [0 Very Low – 21 Very high]
6. Frustration – How insecure, discouraged, irritated, stressed, and annoyed were you? [0 Very Low – 21 Very high]

While the breadth of the type of load measured by this tool provides a lot of value, the six questions can also place their own load on the user when they are asked repeatedly. Many of the future studies in this proposal will have 36 separate tasks, which means that the NASA-TLX questions could be asked and answered 36 times over the course of a few hours. This raises questions about the effects of fatigue and boredom on the participant and calls into question the validity of the results after so much repetition. This led to the initial selection of the Single Ease Questions (SEQ) in the prior study, and it continues to present a challenge in the following studies. However, it may be possible to pair the SEQ with one of the scales from NASA-TLX to test cognitive load – either the Mental Demand or Effort scales may suffice.

The Single Ease Question was developed and tested by Jeff Sauro, and is a simple 7-point scale that asks the participant to rate the relative difficulty of the task:

1. Overall, how difficult or easy did you find this task? [1 Very difficult – 7 Very easy]

The SEQ has been shown to perform well – often as well as much more complex tools – and the lightweight nature of the questions makes it a suitable choice for the repetition required of the studies in this proposal [75]. In the prior study, the SEQ was used to show significant differences between the gloves and the difficulty they imposed on the tasks. The SEQ is also a versatile bridge between the quantitative and qualitative performance data, as it can be recorded easily in both datasets, and used to contract the evaluations between them.

Finally, the Paas subjective rating scale has been shown to be effective in evaluating cognitive load [65]. The tool poses a 9-point scale to the user in the form of the following question:

1. In solving or studying the preceding problem, I invested: [1 very very low mental effort – 9 very very high mental effort]

This scale has been shown to be effective to differentiate between perceived periods of low and high cognitive load and should be effective in pairing with the SEQ to form a set of task-specific questions.

4.3.2 Scaled Response Questions

In contrast to the previous set of questions, these scaled response questions were developed through the initial pilot period of the glove encumbrance study, as they represented the most common answers provided by subjects in the open-ended interview sections of the pilot. After seeing the high rate of occurrence for these observations, these questions were developed to allow for more precise collection and analysis of what seemed to be potential key findings from the study.

The first set is specific to the typing test, and grew out of the reports from participants as to how their natural typing styles and eye positions were changed by typing with the gloves:

1. I was comfortable typing in this task [1 Strongly disagree – 5 Strongly agree]
2. I was able to type normally in this task [1 Strongly disagree – 5 Strongly agree]
3. I was able to maintain my normal eye position in this task [1 Strongly disagree – 5 Strongly agree]

These proved to show a significant difference between the various gloves, and clearly showed the effects of the adaptation and increased load on the participants. As the typing test continues into later rounds of the study, it seems valuable to maintain these questions as part of the protocol.

The second set of scaled response questions were developed from the pilot for the remaining five tasks and relate the participants overall experiences with the individual gloves. These questions are asked at the conclusion of each set of six tasks with a specific glove, with the exception of the “no glove” condition.

Think about your experience in the previous tasks with these gloves. To what degree do you agree/disagree with the following statements?

1. These gloves were comfortable [1 Strongly disagree – 5 Strongly agree]
2. These gloves fit me well [1 Strongly disagree – 5 Strongly agree]
3. These gloves felt slippery [1 Strongly disagree – 5 Strongly agree]
4. These gloves felt grippy [1 Strongly disagree – 5 Strongly agree]
5. I felt the seams of these gloves [1 Strongly disagree – 5 Strongly agree]
6. The fingers on these gloves got in my way [1 Strongly disagree – 5 Strongly agree]
7. I had difficulty moving my hands in these gloves [1 Strongly disagree – 5 Strongly agree]

These questions again are drawn from the most common comments during the pilot and start to establish the principle themes of the interviews. Fit and comfort were obvious factors that influenced the performance of the various gloves, and it is useful to measure the participant’s perception of these effects. Material properties of the gloves that led to slippery or grippy interactions with objects were also very commonly discussed, along with the interference of seams, and general hand movement. This set may be updated with successive study rounds, as new questions emerge from the pilots.

4.3.3 Preference questions

Following the completion of the task sets, the final portion of the study focuses on the participant’s preference – asking them to consider their overall impressions of the gloves, and select the one they would most prefer to complete each task with.

1. Which glove was the most comfortable to wear?
2. Which glove fit you the best?

3. Which glove best worked for you to complete the [x] task? (asked for each task)
4. Which glove best worked for you overall to complete these tasks?

These results are useful to assess the general sentiment users had for the performance of each glove and yielded some unexpected and valuable results in the previous study.

4.3.4 Semi-structured interview questions

The final set of questions follows a semi-structured interview format, to allow for free responses from the participant. They are prompted with the following questions, and the facilitator should feel free to follow up on any contributions offered by the participant.

1. Why do you feel this glove best worked for you?
2. What features of the gloves had the most effect on your ability to complete these tasks?
3. Did you experience any discomfort wearing these gloves?
4. Is there anything about your experience today that we didn't talk about?

4.4 Physiological Measures

The study will ask participants to take a brief set of physiological measures. The first is a set of 4 dimensions, taken from the hand through palpation and landmarking with a set of callipers.

- Distal phalange length [mm]
- Distal phalange width [mm]
- Finger length [mm]
- Palm Length [mm]

These measures will help show any connection between perceptions of fit and task performance in future studies.

The second set of measures concerns the dominant hand of the participant, which has been used to determine the handedness of previous tasks.

- Left
- Right
- Neither

4.5 Demographic Measures

To understand the basic demographic distribution of the populations in this study, these descriptors will be requested:

- Age
- Gender identity

4.6 Inclusion Criteria

This study will recruit up to 100 participants with normal ability to grasp or touch objects with their hands, normal auditory and tactile sensory ability, and no known neuropsychological condition (verified on consent form) for a maximum of 5 sessions lasting up to 180 minutes each. The sample will not seek to balance gender or recruit participants with specific expertise (e.g., those with a particular skill such as in sports).

4.7 Protocol

The protocol for the future studies follows closely with the Glove Encumbrance Study protocol, while implementing structural changes highlighted earlier.

4.7.1 Ordering effects and sample size

The protocol strategy to deal with ordering effects is to employ a Latin square sequence to structure the studies. This approach takes the total number of experimental conditions and creates a structure that ensures that each condition is equally likely to be tested in each position in the order. This helps ensure that the order of the tests does not affect the performance outcomes of the study. This counterbalancing strategy also places the baseline condition in the Latin square set, meaning that it follows the same ordering rules as the other conditions.

This approach ensures that ordering effects are controlled for the glove order, but it is also necessary to address orderings effects for the tasks. This is accomplished by randomizing the task order, using the randomization system built into Qualtrics – the survey software being used to collect the mixed-methods data for these studies. In practice, the facilitator will follow the pre-set ordering document to establish the glove order for each participant, and then follow the generated order from the Qualtrics app to determine the task order. In this way, ordering effects will be effectively counterbalanced and controlled.

4.7.2 Protocol schedule

This schedule is designed around task sets, where each testing condition gets a set. For six testing conditions, there should be ~3 hours of tasks, with 40 additional minutes allotted for consent, training, breaks, and wrap-up. If desired, the session can be broken into two sessions on separate days.

1. Welcome & Consent (10 minutes)
2. Task training (10 minutes)
3. Run task sets
4. For each task set (~20 minutes each)
    a. Don gloves
5. For each task:
    a. Run task
    b. Gather performance data
    c. Ask post-task questions
6. Following each task set (~10 minutes each)
    a. Post-set questions
    b. Post-set interview
    c. Doff gloves
7. Break after half of total sets: (10 minutes)
8. Run second half of task sets, as before
9. Following final task set (10 minutes)
    a. Post-session measures
    b. Post-session questions
    c. Feedback and wrap-up

4.8 Study Evaluation

This is a mixed methods study – balancing a quantitative approach and a qualitative approach – so the evaluation of the data will follow the best practices for each [17].

4.8.1 *Data analysis*

For the quantifiable performance data, a standard Analysis of Variance (ANOVA) test can be performed to compare the means of each condition for the presence of an effect. The quantitative data collected is non-parametric, which indicates the Kruskal-Wallace test as the best fit for this analysis. These tests, run in SPSS, will deliver a hypothesis test summary, pairwise comparisons between each glove condition, frequency charts, and box plots of the data – all of which should enable further analysis.

For the qualitative data, a similar approach using the Kruskal-Wallace test is also useful for the evaluation of scaled-response questions [76]. The results of these tests can also be rendered in a similar manner to the quantitative data. The facilitator's notes will capture experiential results and can be analyzed using various qualitative methods. The themes that emerge from this analysis will help flesh out the story being told by the more controlled scaled response data and show any user perceptions and experiences not captured by other means.

4.8.2 *Hypotheses validation*

These tests can be used to first determine if the glove conditions have any significant differences between them, and then be used to test any of the remaining hypotheses. The hypothesis test summary will determine if any of the null hypotheses are proven out, and the pairwise comparisons should prove out whether any of the remaining hypotheses can be proven with significant results. At this point, the quantitative and qualitative data can be compared to see if they correlate within the same conditions.

CHAPTER 5. DESIGN OF RESEARCH APPARATUS

5.1 Introduction

The final two studies of this project were planned, piloted, and run during the summer of 2020, during the COVID-19 pandemic. As companies and schools responded to the emerging crisis, the team planning these studies had to adapt to new Work From Home (WFH) policies, the shutdown of lab space, and a moratorium on in-person human subjects research. Operating a study under these restrictions required a series of design processes, as new apparatus needed to be developed to support the research. This chapter details the development of the Independent Variable Gloves, the Tactile Discrimination Test, and the Remote Protocol.

5.2 Independent Variable Gloves

5.2.1 Background

The third study in the Cost of Haptics project series was designated the "Constructed Glove" Study, as it was differentiated from the earlier studies which used off-the-shelf gloves. While these yielded results that pointed to significant differences between the gloves, it was challenging to determine if specific design features were responsible for that variance, as they were aggregated together in the various gloves. As this research aims to develop a testing protocol to evaluate the performance of different design alternatives early in the smart glove design cycle, it was clear there needed to be a mechanism to evaluate the effects of individual design features. This necessitated testing gloves that only vary by a single independent variable, which can then be associated with any significant variance in performance. To do this required the construction of a custom set of gloves, each built to a design specification that tests these independent variables.

The results from prior studies informed the requirements for this specification, as the designs of gloves were based off the results from the first study. Previous studies showed a strong preference by users for the Powermesh gloves, which suggested that a similar design be adopted as the baseline glove for this study. On top of this base, the previous evaluation of slipperiness/grippyness of the gloves was explored producing design variations that test various surface finish conditions. Additionally, a variant was produced that featured vibrotactors (VT) placed over the fingertips, which simulates the experience of a basic haptic glove.

5.2.2 Design Problem

This research therefore presented a design problem that needed to be solved – a set of gloves must be constructed to allow for the evaluation of these various design variations. As the requirements called for novel design characteristics in the gloves, a design process was undertaken to determine the specifications for this set of gloves, and to identify the optimal materials and fabrication processes to construct them. This process was undertaken in collaboration with a Functional Garments production team, who were ultimately responsible for the fabrication of the gloves.

To meet the needs of this research, it was determined that five design variants were required:

- Glove variant 1: Basic powermesh glove
- Glove variant 2: Basic glove built from a stretchy knit (Nilo)
- Glove variant 3: Basic powermesh glove + 50% coverage palmar TPU
- Glove variant 4: Basic powermesh glove + 100% coverage palmar TPU
- Glove variant 5: Basic powermesh glove + embedded VT at fingertips

These five variants allow for testing of a set of hypotheses that are each linked to the single variation in the design between the Glove 1 baseline and the single variation in the other gloves. Glove 1 (baseline) replicates the design features of the Powermesh glove used in the Commercial Glove study, while Glove 2 (Nilo) approximates the surface finish of the Running Glove. However, as both gloves are built using the same pattern and construction technique, they can be more accurately compared as they only vary in the material used in their construction.

Figure 35 – The six glove conditions for Study 3: Constructed Glove Study

Glove 4 (100% TPU) attempts to replicate some of the material properties of the Gardening Glove, which featured a tacky Nitrile dip finish that participants reported giving them superior grip. By laminating a TPU film onto the palmar surface of the glove, it was possible to give the Powermesh glove these same grippy characteristics with a single design variation. As a means of introducing a new testable hypothesis, Glove 3 (50% TPU) was introduced to test if a perforated TPU film laminate on the palm would allow for both the greater sensory transparency of the Powermesh with the enhanced grip of the TPU.

Finally, Glove 5 (VT) was specified as a basic Powermesh glove with the addition of pancake motor vibrotactors placed over the hemi-pulp of each fingertip, as this is a basic design that prioritizes haptic feedback for VT gloves, and is common in industry. By building a means of fixing these VTs onto the baseline glove, it was possible to test the encumbrance effects introduced by these motors without other confounding effects.

Each of these design variants allows for direct comparison back to the baseline glove and a no glove condition, which enabled the focused data collection required by Study 3. This is possible as all test conditions shared the same control condition – the participant not wearing a glove – and the same null hypothesis. This allowed for a streamlined protocol that can yield data to test multiple parallel hypotheses.

5.2.3 Requirements

The most basic requirement for this set of gloves was that they need to be constructed in a way that controlled design variables to ensure that only one feature was changing per glove variant. This required a well-documented design process with a high degree of precision and necessitated recruiting the Functional Garments team in the lab.

To position this study in a real-world context, this team was requested to structure the design process for the gloves as they would on a typical project. This was done to reduce the complexity of the fabrication pipeline, as they would work with largely familiar materials and processes. Following the lead of Functional Garments team lead also allowed the Research team to focus on the independent variable design features of the glove and allowed the fabrication experts to manage and control the rest of the design. This relationship mirrors the likely roles that would be played by an expanded team applying the Cost of Haptics evaluation framework.

The Functional Garment team therefore made recommendations for a base pattern, lamination process, fabrication process, and other technical elements. The pattern they recommended needed to be suitable for each of the design variants and needed to be based on design that was not bound by protected IP – suitable for publishing. The pattern also needed to fit the general population as the study did not exclude recruitment based on hand size. The number of sizes needed would determine the total count of gloves needed to be produced for the project, in each of the five design variants.

Finally, the COVID-19 pandemic produced an unexpected set of requirements – the design and fabrication of these gloves had to be conducted by a team working in remote conditions outside of the lab, with whatever equipment was available in their home studios. Design validation would need to be done in an ad hoc manner with team shipping small samples under controlled conditions between their homes. Access to the lab would be delayed for the entire duration of the design and fabrication period, so alternative methods would need to be devised to carry out the build process.

5.2.4 Base pattern

The base pattern that was ultimately selected was based off the specification of the User Study Glove – a parametric glove pattern that had been designed by the Functional Garments team to support studies with similar characteristics to the one detailed here. This pattern supported fast prototyping with a high degree of repeatability and could produce variants easily with common fabrication tools. For this project, the patterns would be finalized in a vector drawing tool, cut with a laser cutter, laminated with a heat press, and assembled using an industrial sewing machine.

The pattern was designed for a snug fit with a stretchy mesh material, to allow for stable positioning of components onto the glove substrate, and to limit excess fabric bunching. Each glove offered full coverage of the hand, terminating with a wide cuff below the wrist. The parameterized pattern allowed for trackable fit concessions and for the integration of additional design features on top of the base pattern. This feature was used extensively to position the TPU laminate and VT pockets, and ensured that they were placed accurately across the range of sizes.

Sizing for the User Study Glove was determined based on an internal study of hand measurements, and was graded for seven standard sizes, ranging from XXS to XXL. After consultation with the Functional Garment team, it was determined that a range of five sizes would be sufficient to cover the general population requirements, given that both fabrics used in the construction of these gloves featured a highly elastic two-way stretch. Each of the selected sizes necessitated the construction of a pair of gloves for each of the five design variant conditions, which produced a set of ten gloves at each size level. Given the logistical complexity of the remote research protocol, it was determined that two full sets of each size/variant combination should be constructed. This ultimately determined the final count for the order, which set at 100 total custom gloves needing to be fabricated.

5.2.5 Laminate exploration

Before determining the final design for the order, an exploratory design process was undertaken to determine the specification for the TPU material, the patterning of that material against the underlying glove, and the fabrication process for laminating the final cut parts to the fabric substrate. This started with a feasibility review of what materials were available during the initial period of the pandemic shutdown, and a short list was created.

Figure 36 – Photos showing the lamination test coupons, with details of the test pattern variations.

To evaluate these materials, a series of evaluation coupons were created, each featuring a different combination of fabric and TPU, with variations in cut pattern and the number of laminated layers. Three TPU candidates were selected – Bemis 3405, Bemis 3415, and Framis 8214 – each an elastomeric sew-free tape with subtly different properties. These were selected as they were an appropriate fit for the heat-press lamination methods specified by the fabrication team, and they would provide varying degrees of grip when applied to the surface of a glove.

Six coupons were produced in total – three Powermesh swatches and three Nilo swatches, each laminated with one of the TPU candidates. The coupons were passed between members of the team for review, and were evaluated for durability, fabrication performance, and their suitability for the requirements. Selection of the material at this stage set the baseline characteristics across each of the TPU gloves, so it was important to understand what the experience of the final glove might be like in each combination.

To accomplish this, an ad hoc pinch test was devised where the test patch of each variation on each coupon was used to cover part of the finger while picking up a small plastic object. Each coupon had six test patches – a solid field, a chevron pattern, and a dot pattern, each in a single layer or doubled sandwiched layer of the TPU. This task simulated many of the core interaction characteristics of the task set in the full protocol and allowed for comparison between the different materials.

The initial evaluation was focused on the grippy feel of each TPU. Placing the solid field of the single layer of the thermoplastic against the test object while picking it up with a pinch grasp over the coupon allowed for a quick test of the TPU's grip against smooth surfaces. It was quickly evident that the Bemis 3405 offered the best additional grip, as the other two materials slid more easily against the object surface. This may be due to the soft hand and increased elastomeric properties of the Bemis 3405, as it conformed more readily to the object surface.

The next evaluation focused on the patterns – each presenting an option for the 50% coverage TPU glove. Across all three materials, there were notable differences between the chevron pattern and the dot pattern. The chevron pattern was relatively shallow and seemed to promote an increased slip of the object running along the rails of the material. The dot pattern, in contrast, seemed to provide a good amount of control regardless of the position or angle of the pattern in reference to the object, which quickly led to its selection for the final design.

The last characteristic for evaluation was the choice between a single layer or dual layer lamination of each TPU. The dual layer arrangement featured a sandwich of TPU on either side of the fabric layer, which significantly increased the strength of the heat-press bond. This enhancement was due to the open knit nature of the powermesh, which only thermally bonded to the TPU where fibers made direct contact. A double layer adhered successfully to itself through the mesh, which made for a more durable construction.

However, the dual layer design led to an inferior user experience in two areas – sensory transparency and skin contact. The thicker layers of the doubled TPU inhibited sensing of object features through the material, which was certain to increase the encumbrance of the gloves in any tasks that required haptic transparency. The interior layer of TPU also stuck readily to the tester's fingers, which decreased adhesion to the task object – reducing the overall grip capability of a glove built with this design. This suggested that a single layer of TPU provided an enhanced user experience that better matched the requirements for the gloves. To validate that the single layer process would be durable enough for fabrication, an agitation test was done by placing the coupons in a warm dryer.

Following this ad hoc evaluation, it was determined that the most optimized combination of material and pattern was a single layer of Bemis 3405, placed with the TPU layer on the outside of the palmar layer of each glove, utilizing the dot pattern to render the 50% coverage variation. While this evaluation was less rigorous than what would be possible in the lab, it allowed the design team to work through a compressed set of design issues during a period of remote work, and led to the successful selection of a set of constraints that allowed the design to move forward.

5.2.6 Pattern refinement and verification

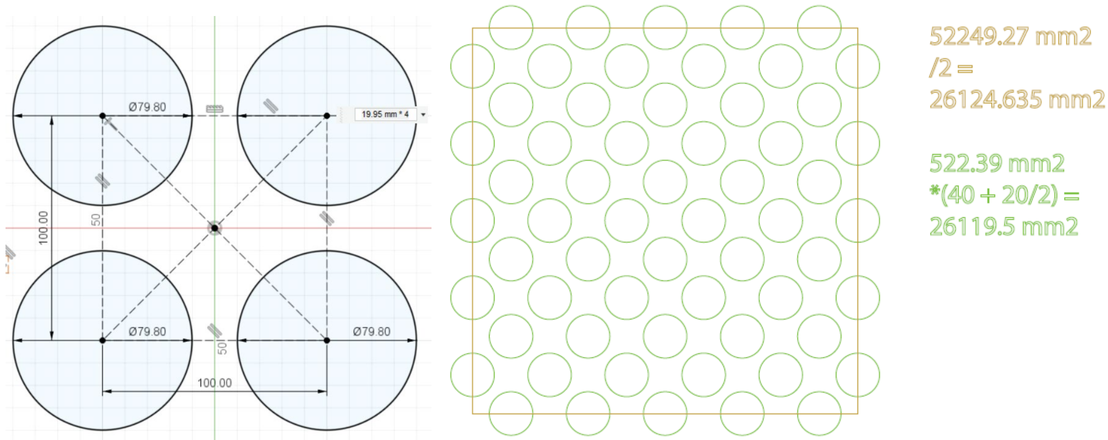


Figure 37 – Detail schematic of the layout for the 50% coverage dot pattern.

Following the selection of the materials, the team worked to define the final elements of the pattern. The first task was to create the 50% coverage pattern for the Glove 3 variant. The hypothesis attached to this variant asserts that dividing the surface area of the palmar face equally between TPU coverage and the basic mesh will yield the user some advantage in providing both increased grip, while maintaining sensory transparency. This requires the pattern to accurately control the dimensions of the core features, to ensure the final pattern hits this 50% coverage target. After working out the dimensional relationships, a parametric pattern was developed in Fusion 360, and swatch was tested for accuracy. After confirming the 50% coverage target had been reached with the tiled pattern, a TPU cut pattern was developed on top of the fabric pattern in the glove model.

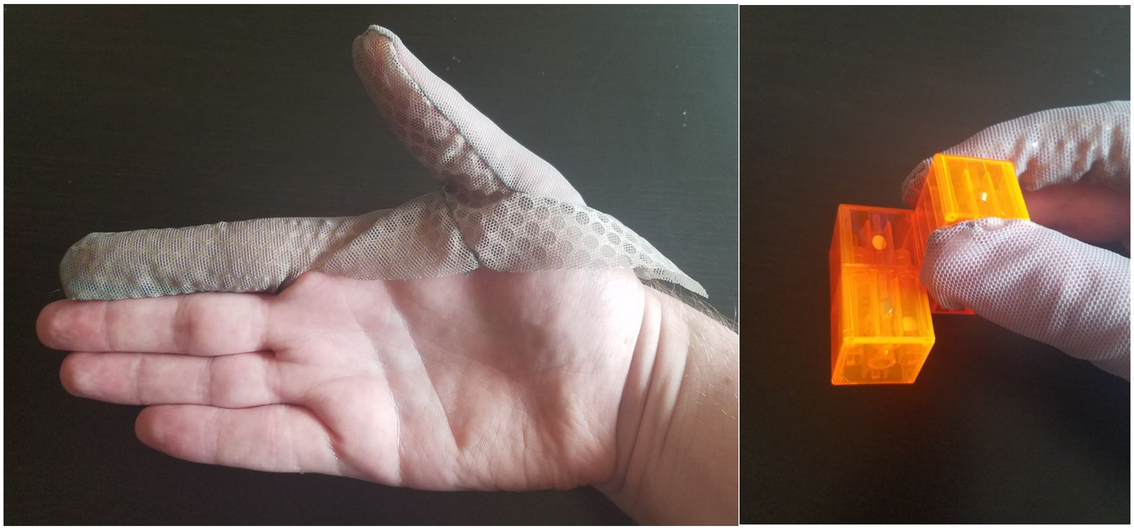


Figure 38 – Photo showing the index/thumb pinch test pattern, and in use for a basic pinch test.

The original design assumed the dots would be positive cut pieces (i.e., the TPU parts), and a series of coupons were produced to test the sizing against the glove pattern, and the lamination process. After performing some abrasion tests, it was clear that the small dots were not making enough contact with the fabric to bond successfully – due to the open mesh pattern on the Powermesh fabric. Successive tests showed that inverting the dot pattern – making the dots the negative space – created a more stable continuous field of TPU that adhered very successfully to the fabric substrate. A test sample was created in the form of a pinch glove – comprising the thumb and index finger – and was tested using the pinch test. This confirmed the stable bond of the TPU to the fabric using this pattern, and a TPU cut pattern was produced.

The first iteration of this cut pattern was designed to provide coverage for the TPU parts into the seam allowance of the glove, with the expectation that sewing through the TPU would help stabilize and secure the lamination. A series of preliminary glove sample were then created for testing. The ad hoc wear tests confirmed that the TPU pattern and lamination process were all working as expected, but the exposed a new problem – abrasion from the cut TPU edges on the seam allowance of the glove fingers. While the TPU is soft to the touch, the laser cut edges has fused a harder edge onto the material, which irritated the fingers of the testers when they donned the sample gloves. This irritation would draw excess attention for the study participants and interfere with their ability to interact with the test apparatus. It was therefore determined that the TPU pattern coverage should be withdrawn from the seam allowance, leaving the softer Powermesh fabric as the only material on the inside the glove.

Fit validation was the final task before proceeding to production. The underlying glove pattern had been through extensive testing and was graded on data gathered from dozens of lab participants. However, testing showed that fingers on a few of the small sizes were longer than expected, and it was determined that a fit concession was required – effectively altering the underlying pattern to make it fit the use-case of the project. These studies are quite sensitive to the fit of the glove at the fingertips, as the gloves are being used in object interaction that require dexterity and sensitivity. As the study was not explicitly testing glove fit, it was important to eliminate the potential confound that an ill-fitting finger would introduce. The team produced a new set of samples with the concession in place, and the design was validated to be ready for production.

5.2.7 Discussion

Producing this set of 100 custom gloves in the middle of a pandemic was a challenging exercise for the team, but a structured design process and a remote testing and validation system allowed for the eventual specification and fabrication of the final design. The team benefitted from extensive prior experience manufacturing gloves and was able to adapt to the work-from-home conditions. However, the nature of the study imposed new restrictions on the design process.

Each of the design features discussed here – the material selection, lamination strategy, TPU pattern refinement, fit concession – could potentially act as a confounding variable for the study if they were not controlled for in the design specification. This set of gloves was built to target a high standard of quality, to ensure the gloves were durable enough to survive the rigors of the study – the TPU delaminating would have interrupted the study, for example. Yet other features, such as the TPU seam allowance, had to be addressed as they could create confounding effects that would interfere with the data collection of the study. Some of the effects that may be observed which show differences between these gloves may be subtle. Therefore, if the participant experiences discomfort or irritation in a way that is not linked to the independent variable being tested, the irritation will interfere with the underlying effects that are intended to be observed.

The independent variables between the glove conditions are simply described, yet each are the amalgamation of numerous design decisions – each of which can affect the glove wearer's experience. As these studies are explicitly measuring the performance of glove design features, care must be taken to ensure the gloves are exhibiting the primary feature that changes between the variants, and not introducing other confounding effects that can overwhelm the detection of the primary variance. This requirement therefore necessitates the introduction of scientific controls and validation into the design process.

5.3 Tactile Discrimination Test

The protocol for this research has tests that collectively measure encumbrance effects on both tasks featuring manual dexterity, and tasks focused on various aspects of haptic sensing. The latter group has primarily featured blinded object identification tasks, such as the Geometric Solids task, and pressure threshold detection tasks, such as the Semmes-Weinstein monofilament test. Each of these prior tasks had limitations that were exposed in the earlier iterations of the protocol, and a project was started to develop a new sensing task with improved reliability and external validity. With the onset of the COVID-19 pandemic, it also became necessary to design a mechanism for delivering a blinded task remotely.

5.3.1 Background

Figure 39 – The apparatus of the STI2 Shape-Texture-Identification Test

The Shape-Texture-Identification Test is an instrument for testing haptic gnosis, or object feature recognition through touch [46]. In this task the subject is blinded from the task object and is shown a reference object, seen in Figure 39 with three shape features and three dot features. The subject is then asked to feel a task object, from which they are blinded, and asked to indicate which reference objects features they believe they are feeling. The test is completed over three difference scales of the shape features, and three different gaps between the dot features, and points are awarded for successful feature identification. This test therefore confirms which features participants can successfully feel and identify across various glove conditions and presents a more controlled and valid upgrade to the Geometric Solids identification task from earlier tests.

As the STI2 test targets patients with neurological injuries, as well as those the who are recovering from injuries to the median or ulnar nerve, the scale of the shapes and pattern features were selected to be sensitive in the measurement range that maps to those conditions. After internal testing, it was quickly determined that none of the gloves in any of the studies induced encumbrance effects on the users that impaired them in this range. That was observed as every participant was 100% correct in identifying the shapes at all scales. This seemed caused by the limited range of shapes – square, circle, and hexagon – and by the size of the lowest scale – 5mm. It was therefore determined that the test would need to adjust these variables to better fit the effective range of the glove encumbrance effects being targeted.

5.3.2 Scaled shapes variation

To meet this new requirement, a variation off the STI2 test was designed with a finer granularity of smaller shape sizes. This granularity was enhanced in two ways – by doubling the count of shapes tested on each plate, and by tripling the total number of plates. The STI2 test has three shapes per plate – circle, square, and hexagon – with a total of three plates with shape diameters of 5mm, 10mm, and 15mm. The revised test featured 6 shapes – circle, square, hexagon, octagon, decagon, and dodecagon. This structure preserved the initial 3 test shapes from the STI2 test, but added three additional conditions, each increasing by a side count of 2. This provided each plate with a series of shapes with gradually changing feature size, as detailed in Table 4. This modification greatly increases the precision of the measurement – moving from 9 possible conditions in the STI2 test to a potential 54 conditions in the modified test.

Table 4 – The ratio between shape side count and shape side length, given a common shape radius of 10mm.

Shape side count	0	4	6	8	10	12
Shape side length (in mm, r = 10mm)	n/a	20	11.54	8.28	6.50	5.36

The updated range of feature sizes also shifts to a range that is likely to create mis-identification results for the shapes in the general population, and therefore is better suited to showing difference in the encumbrance effects posed by the tested glove

This design was then scaled to a set of 9 plates, ranging from 2mm to 10mm in shape diameter, and incrementing in 1mm steps. 10mm is the midpoint of the STI2 test, so this structure allowed continued comparison to results from the predecessor test. 2mm on the lower end was selected as it was the smallest size that could be easily fabricated, and in informal testing seemed to set a lower bound on what participants could identify.

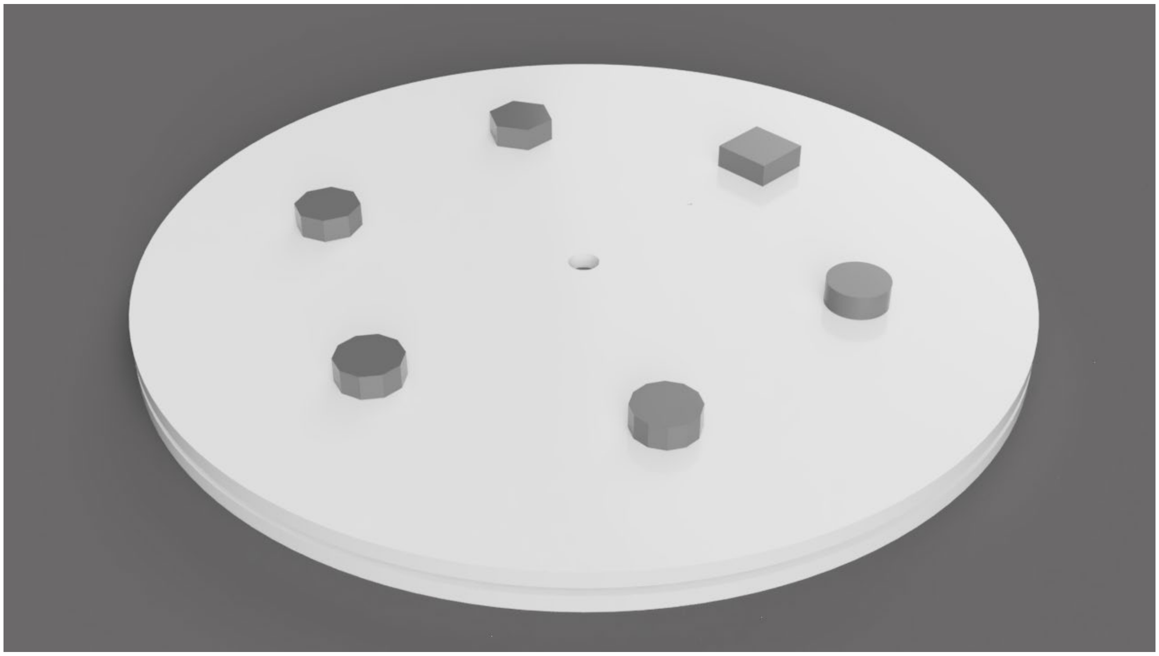


Figure 40 – A rendering of the initial design proposal for the Tactile Discrimination test, feauring 6 geometric shapes of increasing side count.

A design following these updated criteria was then modelled in Autodesk Fusion 360, following parametric modelling procedures. This strategy allowed for easy manipulation of the shape parameters and simplified the process of producing the multiple scaled plates at each size. The plates were designed to be produced using flat sheets of acrylic, cut on a laser cutter, and 3D printed parts to create the shapes. This allowed for rapid iteration of the design, and quickly led to a set of plates which were able to be tested.

The testing proceeded by asking members of the team to identify the number of sides they could feel for a given shape in a blinded informal study. It was readily apparent that the small differences in the features between the shapes – particularly the 8, 10, and 12-sided shapes – were too difficult to distinguish in sizes below 6mm. The modifications to the test therefore were deemed to have increased the difficult too much, and an alternate strategy was devised.

5.3.3 Paired shapes variation

The updated revision to the test proposed to set aside the linear scale of sized elements and focus instead on 3 pairs of test conditions – each testing a different geometric modification. This strategy created more distinguishing characteristics between the shapes, while still offering a pair of shapes that could be easily confused for each other. These changes were made to better calibrate the difficulty of the test – with the goal of making the baseline condition without gloves easier, and the test better tuned to show encumbrance effects from the gloves within the target range of measurement.

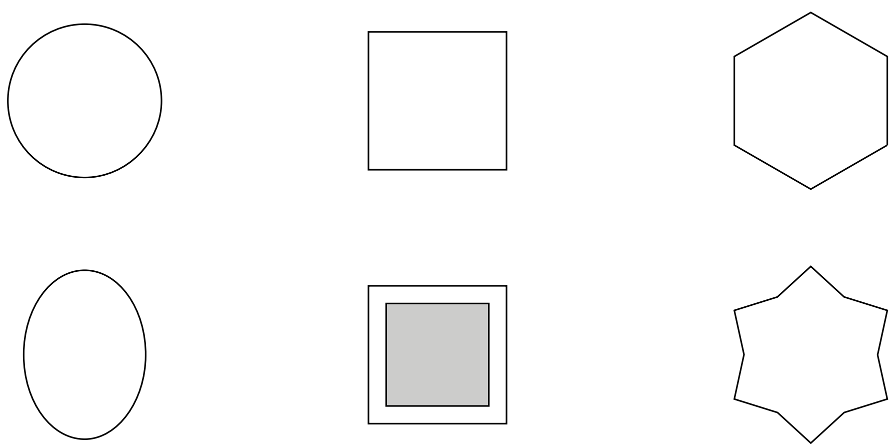


Figure 41 – Illustration showing the six shapes included in the revised Tactile Discrimination test. Clockwise from the top left, these include a circle, square, hexagon, star, hollow square, and oval.

After an iterative design and fabrication process, a new set of tests were created based on 3 new pairs of shapes: a circle and oval, a square and a hollow square, and a hexagon and 6-sided star – seen in Figure 41.

Each of these pairs tested a different relationship between the two shapes. The circle and oval pair tested whether participants could sense the difference in proportion between shapes. The two squares tested whether participants could identify the presence of an indentation in the top face of the shape. The hexagon and star then tested whether participants could feel the presence of an indentation along the edge of the shape.

A new set of plates was fabricated featuring these revised design features, and informal testing was begun. Early results from participants suggested that the new design had found a middle ground in difficulty between the earlier versions, and that participants were able to differentiate the shapes accurately at sizes below 4mm. Running the test with various gloves from the Commercial Glove set led to results in the 8mm range. This seemed to confirm that the revised design was a better match for the encumbrance effects exhibited by the gloves and may be capable of measuring the difference between glove designs. Plans were drawn up to start more formal testing with a pilot study but were interrupted by the COVID-19 pandemic.

5.3.4 Protocol

The protocol for the STI2 test is built around a blinded test of shape identification, and a comparison between the affected hand and unaffected hand of the patient. The patient is shown a reference disk of the STI2 test plates and is asked to identify which of the three shape features they are feeling. The test disk – shown in Figure 42 – is presented to the patient, shielded by a blind so the patients are unable to see the test disks. The test shapes are presented to the patient in a random order, and the patient is asked to identify the shape they believe they are feeling with the labelled identifier next to the shape on the reference disk (A, B, C,..). The patient starts each test with the index finger of their unaffected hand, and then proceeds with a test of the index finger on their affected hand. For each hand, if the patient correctly identifies all three shapes, they then move on to a smaller scale. If they fail to correctly identify all three shapes, the test is concluded for that hand. The final result allows the clinician to compare the patient's hands to each other, and to track their performance over time.

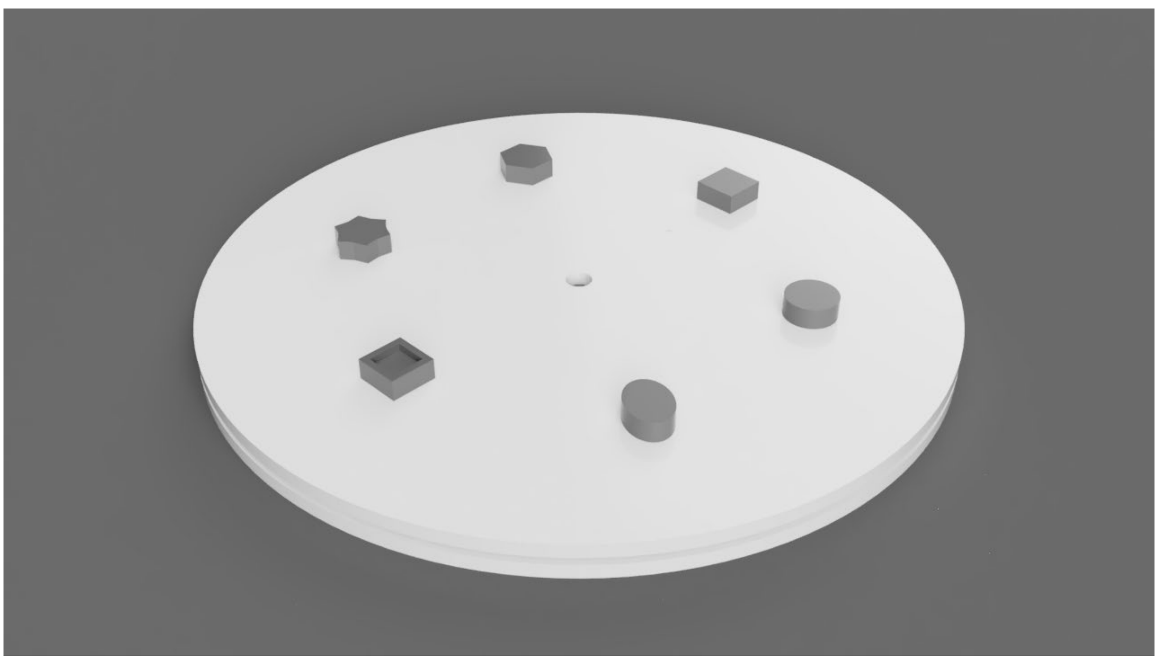


Figure 42 – A rendering of the revised design proposal for the Tactile Discrimination test, featuring 3 pairs of contrasting geometric shapes.

This protocol works well for the target population of the STI2 test and has been validated to be reliable and effective [46]. As previously discussed, this version of the test was too simple to complete for a study population without impaired sensory function in their hand. This was due in part to the design of the protocol – with only three shapes to test, participants in the earlier informal studies reported being able to use process of elimination to help influence their identification of the second and third shape in each test. The STI2 protocol also treats the three shapes on each plate as a single test – all are identified, or the test fails – and it was determined that taking multiple independent measurements would help improve the granularity of the test measurements.

A modified version of the protocol was then pursued, drawing on the common protocol for evaluation of cutaneous sensation level with the Semmes-Weinstein monofilament test [6]. This protocol tests and tracks each test site (commonly, the patient's fingertips) independently, and requires multiple measurements for confirmation. Following this structure, a protocol was devised for the new test, now named the Tactile Discrimination Test, with the goal of identifying the lower threshold for confirmed identification of each shape. This new protocol started the participant at the midpoint disk (5mm) and tested each shape on the disk in a random order multiple times. For each shape condition, the test was continued until the participant correctly identified the shape twice, or failed to identify the shape twice. If the identification was correct, the participant moved on to a smaller scale for that shape and continued at each scale until they failed to correctly identify the shape or completed the 2mm disk. If the identification was incorrect, the test would be repeated at a larger scale for that shape until 2 correct responses were recorded. The final measure was then recorded for each shape of the smallest scale where the participant successfully identified the shape twice.

	Circle (a)			Oval (b)			Hollow Squ. (c)			Star (d)			Hexagon (e)			Square (f)			
Size	A1	A2	A3	B1	B2	B3	C1	C2	C3	D1	D2	D3	E1	E2	E3	F1	F2	F3	Correct
2mm																			0
3mm	x	b	b							e	x	e	x	d	a	c	c		3
4mm	x	x					f	f		x	x		x	x		x	x		8
5mm	x	x		a	x	a	x	x		x	x		x	x		x	x		11
6mm				x	x														2
7mm																			0
8mm																			0
9mm																			0
10mm																			0
Correct		5			3			2			5			5			4		
Smallest	4mm			6mm			5mm			4mm			4mm			4mm			

Figure 43 – An example of the Tactile Discrimination test score sheet, showing the markings for correct answers (x) and incorrect answers at every scale.

An example of the test score sheet is seen in Figure 43, showing how the test progresses through the various conditions. Successful identifications are marked with an x, while unsuccessful identifications are marked with the corresponding letter of the shape the participant misidentified. In this way the shapes are individually tracked, and the scale threshold can be identified.

5.3.5 Fabrication

Fabrication for the Tactile Discrimination test was done by the research team, using the on-campus workshops and labs. Designs for the test apparatus were therefore evaluated against their fit with the available fabrication processes – primarily laser cutting and various 3D printing technologies. The designs underwent many iterations, each testing a different fabrication process. Initial tests focused on producing the entire plate in a single pass, using a 3D printer, which would save post-processing and assembly time. Various models were built and printed using Fused Deposition Modeling, Polyjet photopolymer, and Carbon 3D resin printers. These were each compared for finish and edge fidelity.

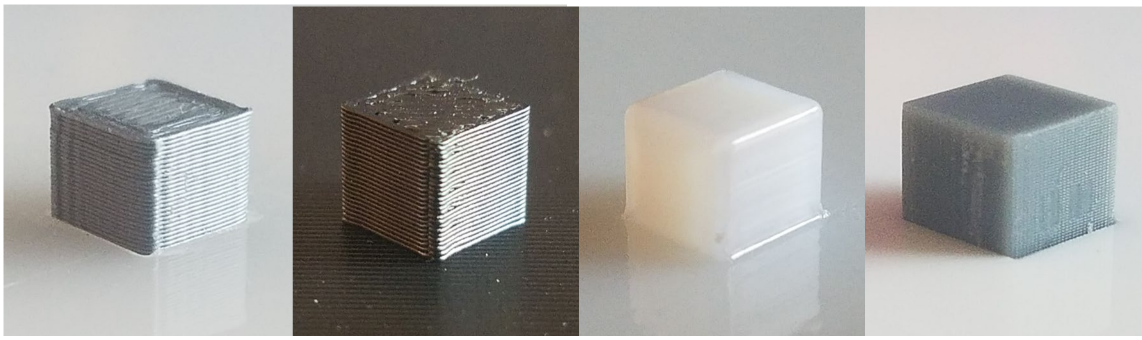


Figure 44 – Detail of the surface finish of the 5mm square shape fabricated using four different printing technologies. (L-R) STi2 FDM precedent, FDM sample, Polyjet sample, Carbon 3D resin print - put into production.

The STI2 test used an FDM process, and so served as a baseline comparison. A seen in Figure 44, the print is clean, with the edge defined the radius of a single printed layer. Prints produced in the lab were able to achieve a similar degree of fidelity, though with a rougher surface finish on the top face. The Polyjet print was seen to be much smoother on the faces, though the finish of the edges had allowed for a slight slump – rounding over the corners and providing a smooth radius along the edges. By contrast, the Carbon 3D resin prints were clean, accurate, and produced crisp corners and edges without layer lines. These were compared in informal testing, and there appeared to be a stark difference in the edge detection between the various prints, with a strong advantage for the resin print. With the priority for this test being the blinded evaluation of shapes, it was determined that the resin print's superior edge rendering made it the best solution.

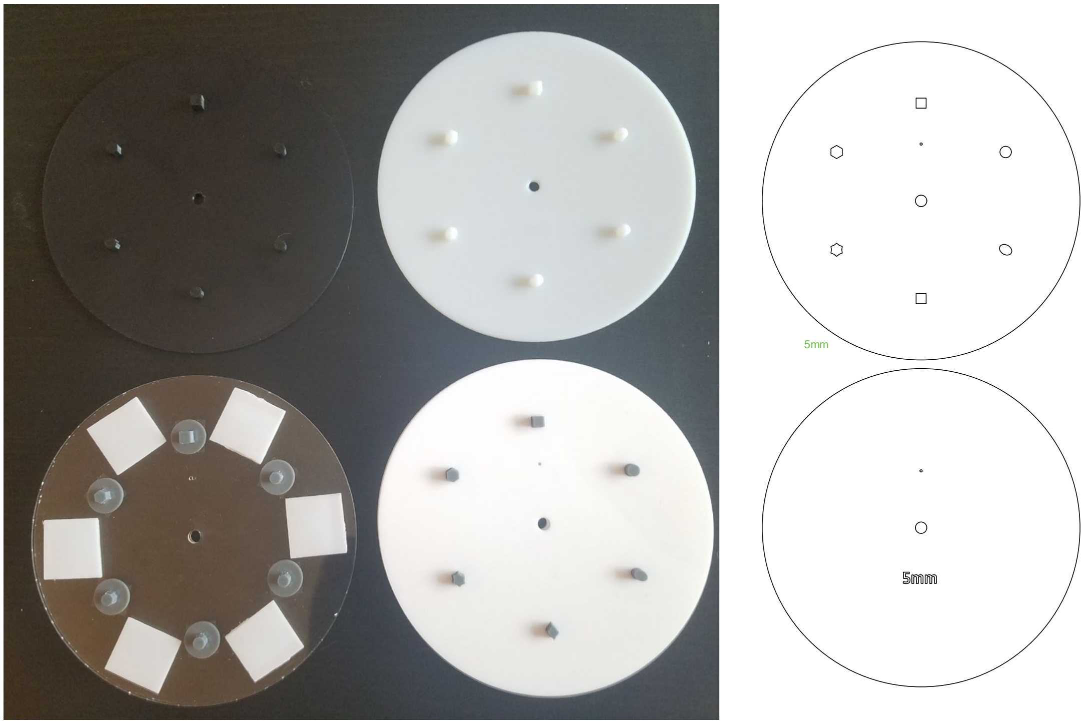


Figure 45 – The iterative design process for the Tactile Discrimination test plates - showing the internal adhesive construction and the laser cutter cut patterns.

The consequence of selecting the Carbon 3D platform as the 3D printer of choice was the limitation on producing a whole test plate in a single process – the build volume on the available printer prevented this. This necessitated returning to a multiple process design, where the shapes were produced on a resin printer, and the plates were produced by laser cutting acrylic sheet. The resulting pieces were then assembled by hand and laminated together with various adhesive tapes to create a finished test plate. The bottom plate was etched with the size of the plate features, and anti-skid feet were added to help stabilize the plate during testing.

5.3.6 Remote Testing Apparatus

The COVID-19 pandemic of 2020 created a new set of challenges for the successful operation of this study. Restrictions on human subjects research in the lab prevented the scheduled pilot from proceeding, and the use of the existing in-lab protocol. Approval was eventually granted for a remote facilitation study, with the apparatus shipped to the participant's homes. This allowed the study to proceed but imposed a unique set of circumstances on the Tactile Discrimination test, due to the blinded nature of the task. Under normal circumstances, the facilitator would select and manipulate the test disks for the participant, both guiding them to the current shape, and managing the test order and randomization. For this task to be included as part of the remote study protocol, a new apparatus had to be constructed to allow the participant to self-administer the test – performing the facilitation function, while remaining blinded to the current test condition.

As the primary means of creating the task order was the rotation of the test disks, it was determined that a mechanism to automate this rotation was required. A system was proposed featuring a turntable platform triggered by a single large button that the participant could press to advance through the different test conditions. The turntable featured a stepper motor to drive the rotating platform, and a microcontroller that could automate the randomization, and deliver a balanced and accurate test. In this system, the button was visible to the participant, while the turntable was hidden behind the blind – allowing the user control over the test. The position of the disk was captured by a camera and conveyed to the facilitator who was be able to run and score the test remotely.

Plans were drawn up to design a system using off-the-shelf parts and a custom laser cut acrylic enclosure – using materials, parts, and fabrication processes that were available during the pandemic. The system was built around an Adafruit Feather nRF52840 Express – a powerful and versatile Arduino-compatible microprocessor breakout board, and a NEMA-17 stepper motor. The rest of the system included all the required power and control components to drive the stepper motor, buttons and switches for the interface, and panel mount jacks to allow external cables to plug safely into the enclosure.

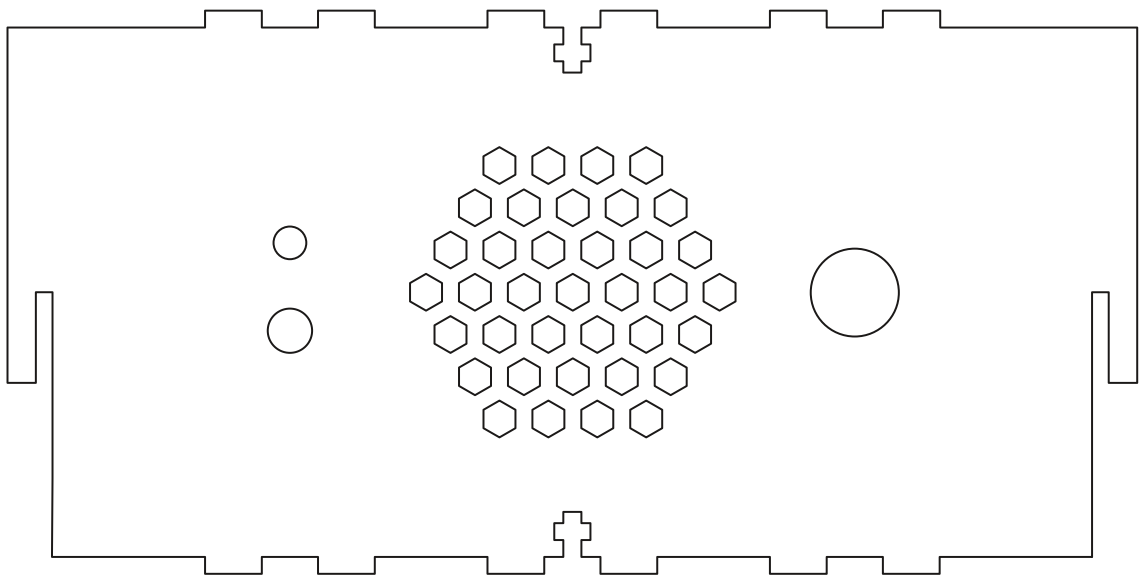


Figure 46 – The snap fit enclosure pattern, showing cutouts for the I/O cables, fasteners, and ventilation.

The enclosure was designed to be built with snap fit acrylic parts and machine bolts, following a technique called Interlocking T-Bolt Construction [56]. This allowed for a complex 3D form to be built out of cut flat panel stock and made it easy to iterate through the design process as individual panels could be modified and swapped out. An example panel design is seen in Figure 46. This approach also meant that the final test plates and the enclosure could be made out the same material, using the same fabrication process, which greatly lowered the complexity of the build during the limited shop availability during COVID-19.

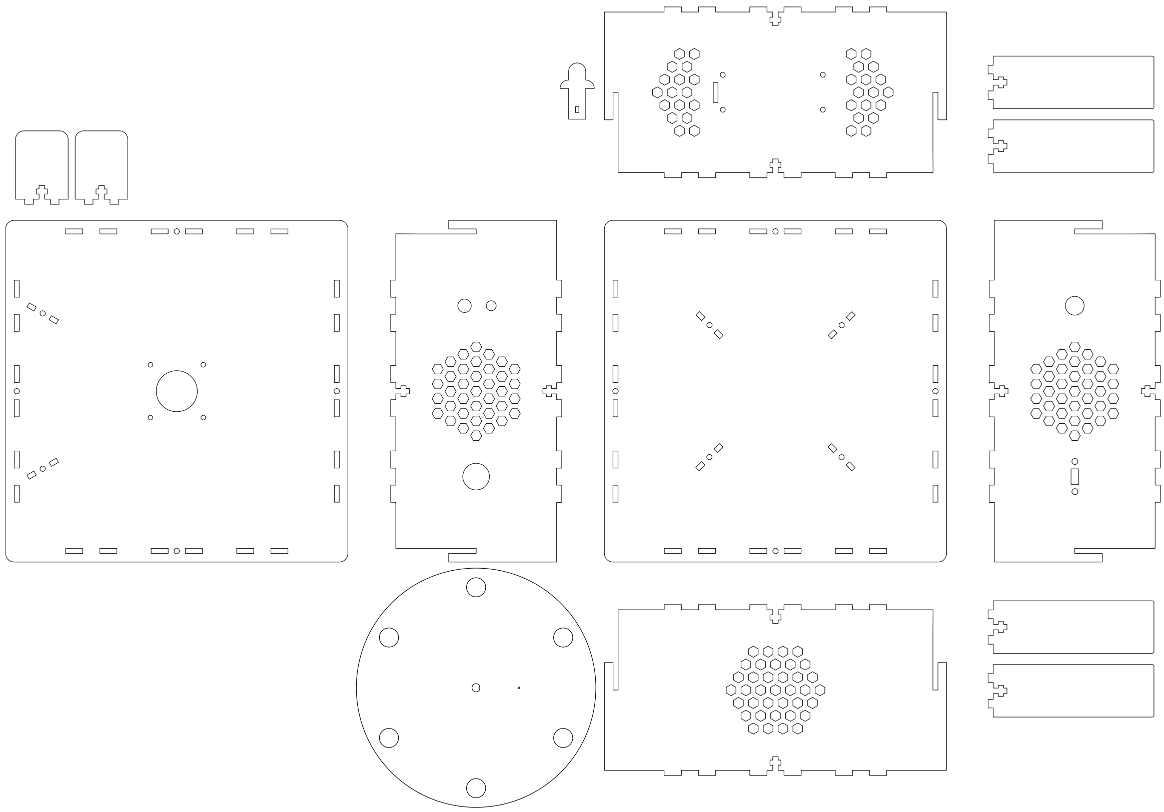


Figure 47 – The complete cut pattern for the Tactile Discrimination test enclosure.

Plans were drawn up for the enclosure in Adobe Illustrator, using the interlocking T-Bolt construction system – as seen in Figure 47. The enclosure was designed to present the turntable on the top face, driven directly by the stepper motor. The 4 faces of the enclosure each presented a variety of pass-through holes for press fit and panel-mounting of the various cables, jacks, buttons, and lights required by the system. Two fins were fastened near the front of the turntable to help guide the participant's fingers towards the correct test shape, and internal supports were added to the bottom face to help support the weight of the stepper system – seen in Figure 48. Venting holes were added to allow for increased airflow over the stepper motor, which produced substantial heat under load. Finally, an acrylic key was cut to act as a power switch, manipulating an internal switch attached to the Arduino.

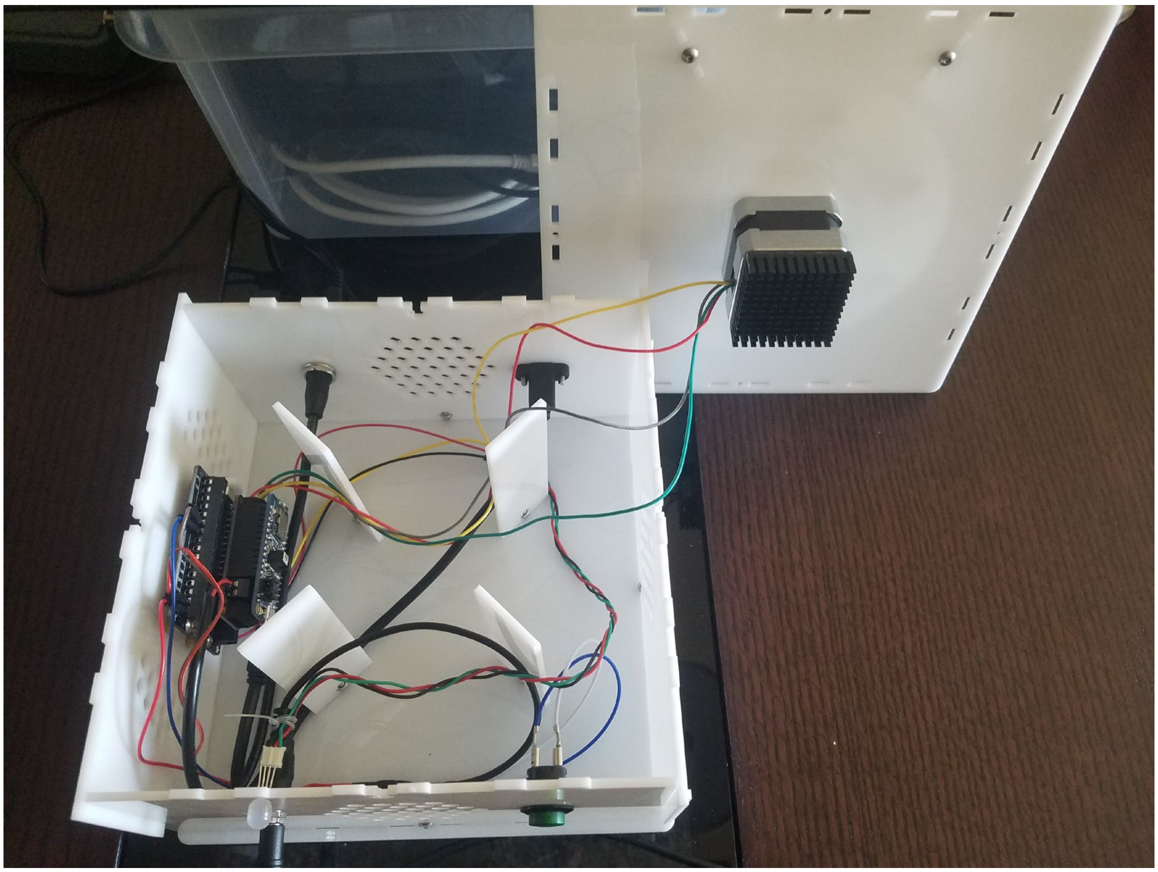


Figure 48 – Internal detail of the Tactile Discrimination test, showing component placement and wiring.

The completed parts were cut and adjusted through a number of revisions, working to make sure that the tolerances of the interlocking parts were tuned to allow for a secure snap-fit. After finalizing the dimensions, a full set of pieces were cut, and assembled together using machine screws. Components were mounted to the case, using the prescribed holes, and wiring was completed to connect the components back to the microcontroller. A ball-bearing turntable was fit to the top of the enclosure, and a turntable plate was fit on top of it, keyed to the stepper motor shaft. This plate fit the anti-slip feet placed on the bottom of the test plates and enabled precise and repeatable control over the swappable test plates.

The final system component was the interface button for the participant. The requirements for this button dictated that it would be capable of controlling the system while the use was focused on the blinded task. This meant the button needed to be large, easy to operate, accurate, and provide feedback to the user. After investigating several options, the promotional Staples Easy Button was selected [79]. This was made possible due to a guide that demonstrated how the button can be easily re-purposed, allowing it to be integrated into any microcontroller system [48]. By cutting a few traces on the internal button PCB and soldering a new cable into the right pads, the button could be read with a normal analog input on the microcontroller. This button's weight, mechanical force-feedback, anti-slip pads and solid construction made it a reliable interface for this remote system.

Figure 49 – The operational Tactile Discrimination test remote apparatus, showing the user-accessible button in the foreground.

The system hardware was completed, seen in Figure 49, and programmed with custom code to drive the task interaction. On boot, the system powered on the stepper motor, and loaded each of the six task positions into an array in a random order. As the participant pressed the button, the system would step through the array, and rotate to each of the 6 shape positions in turn. When the sixth position was reached, the system generated a new random order for the next set of six. In this way, the system ensured that the shapes were presented to the participant an equal number of times, while keeping the order of the presentation randomized. A physical reset button was provided that re-generated and re-started this sequence. Finally, a debugging system was constructed that allowed for control and testing of the system using the serial monitor.

5.3.7 On-screen display

With the hardware developed, there was one additional challenge to solve – supporting the remote facilitator as they guided and scored the task. A camera was specified to be pointed directly at the top plate, and send a live feed for viewing by the facilitator. However, given the small scale of the shapes, the features were deemed to be too small to reliably be identified by sight using the camera. An on-screen display, shown in Figure 50, was therefore developed to augment the camera feed and ensure that the rotation of the plate was obvious to the facilitator.

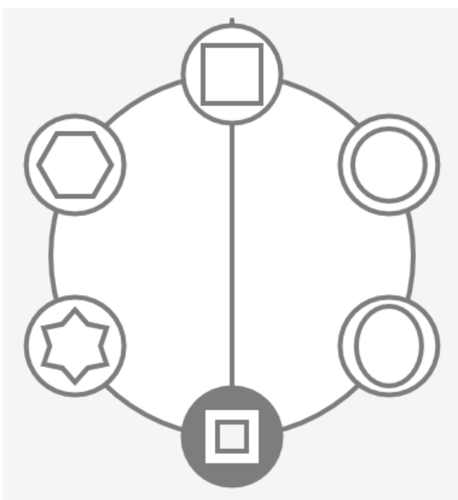


Figure 50 – Detail of the on-screen web interface for Tactile Discrimination test, showing the animated rotation indicator that shows the current state of the test apparatus.

A lightweight app was then developed in Javascript using p5.js, capable of creating a small visualization that ran in a browser window. The app showed a schematic representation of the test plates and rotated along with them whenever the participant pressed the button. This communication was handled between the hardware and browser by using the native

USB keyboard emulation feature of the microcontroller – whenever the participant pressed the button, the hardware acted as a keyboard, and pressed a virtual key – signalling the web app to rotate to a given position. This allowed communication to solely happen over the USB cable, which made the system easier to set up and operate for the participant.

5.3.8 Discussion

The Tactile Discrimination Test (TDT) was designed to measure the encumbrance effects that gloves imposed on their wearer – specifically their normal facility for haptic gnosis. It is important to understand how this ability to understand the world through touch is altered by wearing gloves, to better measure the costs imposed on the user in exchange for the benefits they receive by wearing them.

An important distinction between the goals of the TDT, and the predecessor STI2 test, as the latter's focus on directly evaluating the sensory function of the participant's nervous system. While the TDT uses similar measures – the accurate identification and discrimination of various scales of geometric solids – the TDT is designed the measure the encumbrance effects imposed by the gloves, not the participant's baseline functional condition. The principle hypothesis being tested is whether the gloves in question are significantly different from each other in the effects they impose on the user, not whether the user has a neurological impairment.

The STI2 test was calibrated to measure these impairments reliably, building on prior research to establish an effective range. The TDT may measure glove effects that are much closer to each other in the degree of their effects, or with much more variance – there is a significant range of variability in the sensory occlusion possible designs could impose.

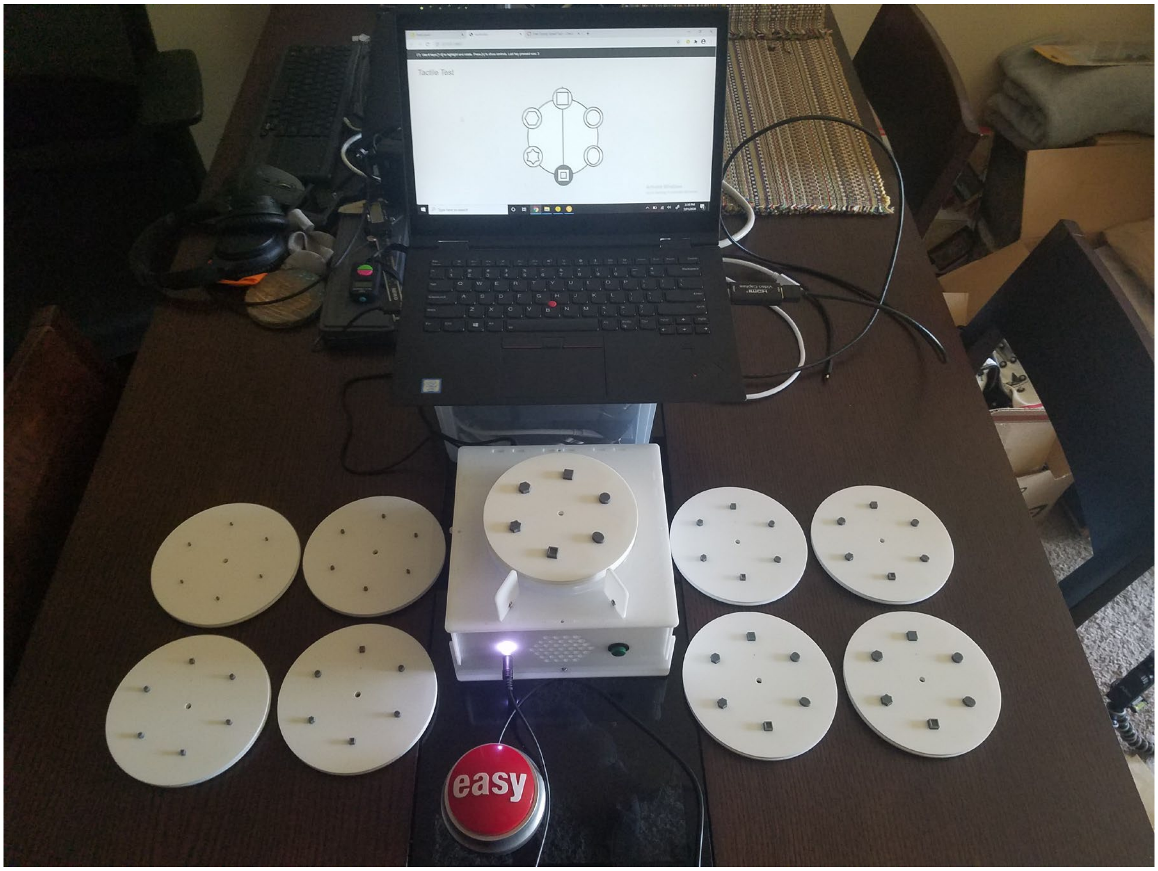


Figure 51 – The complete Tactile Discrimination test remote protocol apparatus.

However, the TDT, shown in Figure 51, should be able to provide some guidance as to the degree of difference between different designs, and a rudimentary comparison between those designs and the baseline (no glove) condition. It is possible that some design features may enhance the participant's ability to discriminate between the shapes – the TDT should also be able to show these outcomes.

5.4 Remote Testing apparatus

5.4.1 COVID-19 Remote Protocol

With the introduction of COVID-19 policies surrounding human subjects research, it became obvious that this study would need to run using a remote protocol. This meant that participants would complete the study in their own homes, with facilitation handled remotely. These conditions therefore required the development of an apparatus that could be packed and shipped repeatedly to participant's homes.

One of the major differences between the remote protocol and a more typical lab operation was the lack of hands-on facilitation. Participants would therefore need to be able to set up the apparatus themselves, manage the transition between tasks, and complete any required sanitation protocols on their own. This is an extra burden placed on the participants, and it was possible that that the added load might lead to fatigue and frustration for them. Given the extra burden – and the mandatory shipping and sanitation waits – it was determined that Study 2 and Study 3 should be run concurrently, as they shared the same apparatus and protocol. This meant that both studies would share the same participant sample, though they would not be considered to be part of the same group across studies.

The existing Cost of Haptics protocol was reviewed, and it was determined that the current task set could be run remotely, with two exceptions. The Semmes-Weinstein Monofilament test requires training to be administered and is not possible to be self-administered. This eliminated it as an option for the remote studies, as there was no practical way to run the test. The Tactile Discrimination test faced similar obstacles, but these were overcome with the development of the remote testing apparatus, as outlined in Chapter 5.3.6. The rest of the tasks are simple enough to set up and can be administered with remote guidance.

It was therefore determined that the studies could be run remotely if solutions could be developed for remote facilitation, data collection, and sanitation. It was also determined that informed consent could be handled remotely, and the modified protocol was approved by Western IRB, and the Georgia Tech IRB (in reliance on the Western IRB approval). A packing plan was developed, and the apparatus was packed and shipped to each participant, allowing the study to move forward.

5.4.2 Remote facilitation

Running these studies remotely required running facilitation using internet tools. A remote study platform called Lookback was selected, which provided a secure mechanism for connecting facilitator and participants though a video call, session recording, and basic notation tools. Lookback provided a mechanism for simultaneous recording of both the participant's desktop and webcam, so it was determined that the desktop could be used to captured additional data for the study.

This first involved providing an additional movable camera feed that could be positioned to capture the test in progress. This allowed the main webcam to be focused on the participant's face, and the moveable camera to focus on their hands in action. This was critical for the success of the study, as all the tests require a clear view of the apparatus by the facilitator to operate and score. To accomplish this, a wide-angle action camera was selected, and mounted to an articulating swing arm with a weighted base. This allowed the participant to quickly reposition the camera to capture the task activity. The video feed from the camera was captured with an HDMI capture card and viewed in a window on screen on the participant's laptop. This allowed the camera feed to be recorded by Lookback's desktop recording feature and provided simultaneous feeds of the participant's face and hands for the recording.

Additional tasks were captured by running their software apparatus in a web browser and recording the results through the Lookback desktop recording function. The typing test took advantage of this by using an online typing test to operate and score the typing task, while the Tactile Discrimination test displayed the web-based on-screen display that was discussed in Chapter 5.3.7. Participants also took the post-task NASA-TLX scale by using an enlarged printed version of the test and pointing to their selected answers on the scale where the camera could capture their choice. Lookback was therefore able to capture all the primary action of the test operation and became the primary data source.

During facilitation, all the tests were scored and recorded using Qualtrics – an online survey tool. Qualtrics provided a streamlined interface to capture the varied data produced by the study and allowed for anonymous recording of the data. The Qualtrics interface also handled the task randomization mechanism, which aided in the counterbalancing plan. Following completion of the study, Qualtrics supported export to an SPSS-compatible format, which aided the start of the data analysis. Facilitation notes were also captured and recorded in a secure drive share. Collectively, these tools allowed for remote facilitation and data collection, and were an essential part of the study apparatus.

5.4.3 Sanitation plan

To best protect the health and safety of the participants during the COVID-19 pandemic, a sanitation plan was devised and approved. This plan utilized three different sanitation strategies: body sanitation, apparatus sanitation, and time sanitation. For the body, participants were asked to follow standard COVID-19 sanitation protocols and wash their hand thoroughly at the beginning and end of every session.

For the apparatus, participants were provided equipment and materials to sanitize both soft and hard surfaces. For hand surfaces, nitrile gloves and 70% alcohol wipes were provided, with instructions to swab every exposed surface with the alcohol solution. For soft surfaces, like the gloves, participants were provided with a wash basin, hand wash laundry soap, and a set of drying racks. Participants were asked to soak the gloves they used in warm soapy water, before gently agitating, rinsing, and drying the gloves. The drying racks allowed the gloves to be dried in a flat orientation, which was determined to be the optimal solution by the glove design team as it would inhibit warping while drying due to gravity.

The final strategy employed was time-based sanitation, which asked the participants to seal the apparatus in boxes and sign a card with the time and date of the sealing. The boxes would then be collected and sent to the next participant, with instructions to delay opening the boxes until 72 hours had passed since the time and date on the card. This provided a backup to physical sanitation strategies, and collectively ensured that the equipment could be safely shipped between participants.

5.4.4 *Discussion*

The remote version of the protocol added many new considerations to the operation of the study and placed many new burdens on both the participants and facilitators. The extra work required of the participants added up to an additional four hours of setup, training, teardown, and sanitation – none of which would be a normal part of participating in a study. They all handled this extra requirement with grace, but it was clear that these burdens introduced extra fatigue, some of which may have influenced the outcomes of some of the tests.

Remote operation also placed a burden on the facilitation team, as they were unable to physically interact with any of the apparatus once the study had started with the first participant. This placed an added burden on the facilitator to guide the participant to set up the apparatus, and to guide them through any of the troubleshooting that needed to take place. This stretched the time to operate each segment of the study out and made for very long days for everyone involved. This fatigue may have also influenced the outcomes of some of the studies.

Despite these additional burdens, this remote protocol allowed the studies to operate during a general moratorium on in-person data collection. This type of study was particularly challenging to adapt, as it involved so many tactile tests, and was developed to be observed in a lab setting. However, the studies were able to proceed remotely, successfully collected all of the required data, and protected the health of participants and facilitators – an outcome that was not guaranteed at the start of the pandemic.

CHAPTER 6. STUDY 2: COMMERCIAL GLOVE STUDY

6.1 Introduction

Following the results of Study 1, a study plan was made to build on the new findings, and transition to the revised task set, outlined above. This next study proposed to again test commercially available gloves that have characteristics that might feature in future glove projects, using the original set of five gloves from the prior study. This provided the opportunity to test the revised protocol and new tasks against known encumbrance conditions, and to compare the results. This also allowed for a final test of the commercial glove set, to be included in the benchmarking effort, as these gloves are not guaranteed to be commercially available in the future. The data collected from this study was intended to confirm the efficacy of the protocol with a different task set.

6.1.1 Hypotheses

The common hypothesis across all conditions were as follows:

> H_0: There will be no significant difference in performance between the glove conditions.
>
> H_1: There will be a significant difference in performance between the glove conditions.

If significant differences in the measured effects between the gloves are observed, a set of secondary hypotheses were then tested:

H_2: The “no glove” condition will have better performance across all tasks than any glove condition.

H_3: The Hockey Glove condition will show significant impairment in performance compared to the “no glove” condition.

H_4: Participants will express an overall preference for the “Basic powermesh glove” condition.

6.2 Methods

6.2.1 Participants

Participants for this study were recruited from a pool of Research Assistants and were approved for participation in a remote study. Due to limitations from COVID-19 policies, the sample was capped at 6 participants (4 self-identified as female, 2 as male), with an age range between 24 and 33 years of age. As discussed in Chapter 5.4.1, the participants for this study were also selected to participate in Study 3, as both studies shared the same apparatus and sanitation protocol. While the same participants completed both studies, they were not considered as single group across both studies, and the studies were not run concurrently. Participants were required to meet the inclusion criteria identified in Chapter 5 and be willing to conduct the study remotely in their own homes. The protocol was approved by Western IRB, and the Georgia Tech IRB (in reliance on the Western IRB approval), and all participants provided their written informed consent.

6.2.2 Study Design

This study was structured to test each combination of glove condition and task, while controlling for ordering effects. This was done by using a Latin square to structure the order of the glove conditions and using a random generator to structure the order of the tasks within each glove condition. Six glove conditions are included in this study, which are discussed in detail in Chapter 3.1.1:

- Glove 0: no glove
- Glove variant 1: User Study Glove - Powermesh
- Glove variant 2: Running glove
- Glove variant 3: Gardening glove
- Glove variant 4: Tactical glove
- Glove variant 5: Hockey glove

6.2.3 Apparatus

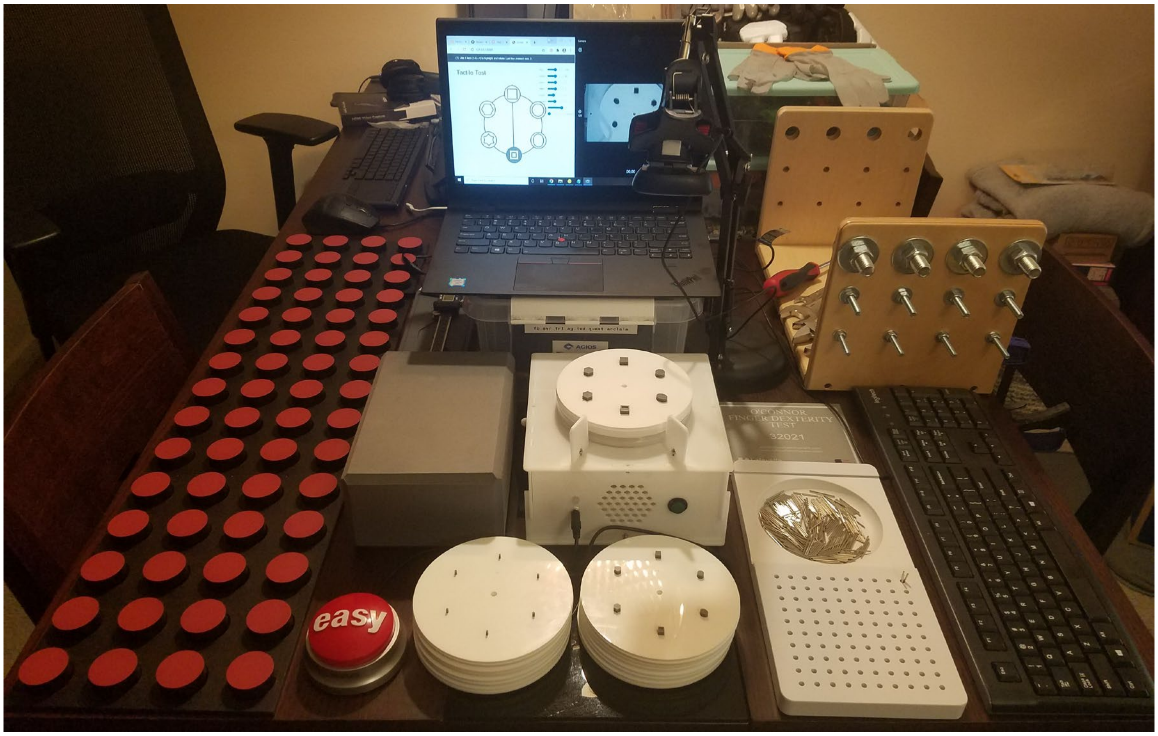


Figure 52 – The assembled apparatus for the Cost of Haptics Study

Five tasks – their apparatus shown in Figure 52 – were included in this study, as detailed in Chapter 4.2:

1. Typing test
2. O'Conner dexterity test
3. Minnesota dexterity test
4. Bennett hand tool dexterity test
5. Tactile Discrimination test

Each of these tasks was administered to the participant using the requisite hardware, which the participant configured under the guidance of the study facilitator. Additional apparatus included a laptop, camera equipment, and cleaning supplies, which were used to support the remote operation of the study, as detailed in Chapter 5.4.

6.2.4 Measures

Each of the five tasks had a primary measure of performance, and a set of associated subjective measures. The performance measures are as follows:

Table 5 – Performance measures for Study 2 tasks

Task	Measure	Unit	Calculation
Typing test	WPM	count	Correct, complete words / minute
	Error rate	%	Correct words / total words
Bennet hand tool test	Speed	seconds	Time to completion
	Error rate	count	Dropped parts during task
O'Conner pin test	Speed	seconds	Time to completion
	Error rate	count	Dropped pins during task
Minnesota dexterity test	Speed	seconds	Time to completion
Tactile discrimination test	Scale	mm	Smallest scale correctly identified

The subjective measures for this study are discussed in depth in Chapter 4.3. The set included for the study are as follows:

- NASA-TLX Scale – Mental load, Physical load, Frustration [0-20]

- Scaled response typing questions [1-5]
- Scaled response glove experience questions [1-5]
- Glove preference questions [choice of preferred glove]

6.2.5 Analysis

Analysis for this study was completed using SPSS for Windows. Performance data for these gloves and scaled response questions are considered to be non-parametric, therefore the Kruskal-Wallace test was selected to determine if there was variance between the median values for each group. The results are rendered in boxplots, which depict the median, quartiles, and outliers. SPSS renders the box as representing the interquartile range (IQ) of the central 50% of the values (ranging from the 1st to 3rd quartile), with the median indicated as a line across the box. The whisker lines show values that extend from the limits of the box to the highest and lowest values that do not exceed 1.5 times the range of the IQ. Circles and crosses indicates outliers, the latter of which represent extreme outliers with values that exceed 3 times the range of the IQ [42].

6.3 Results

6.3.1 Typing test

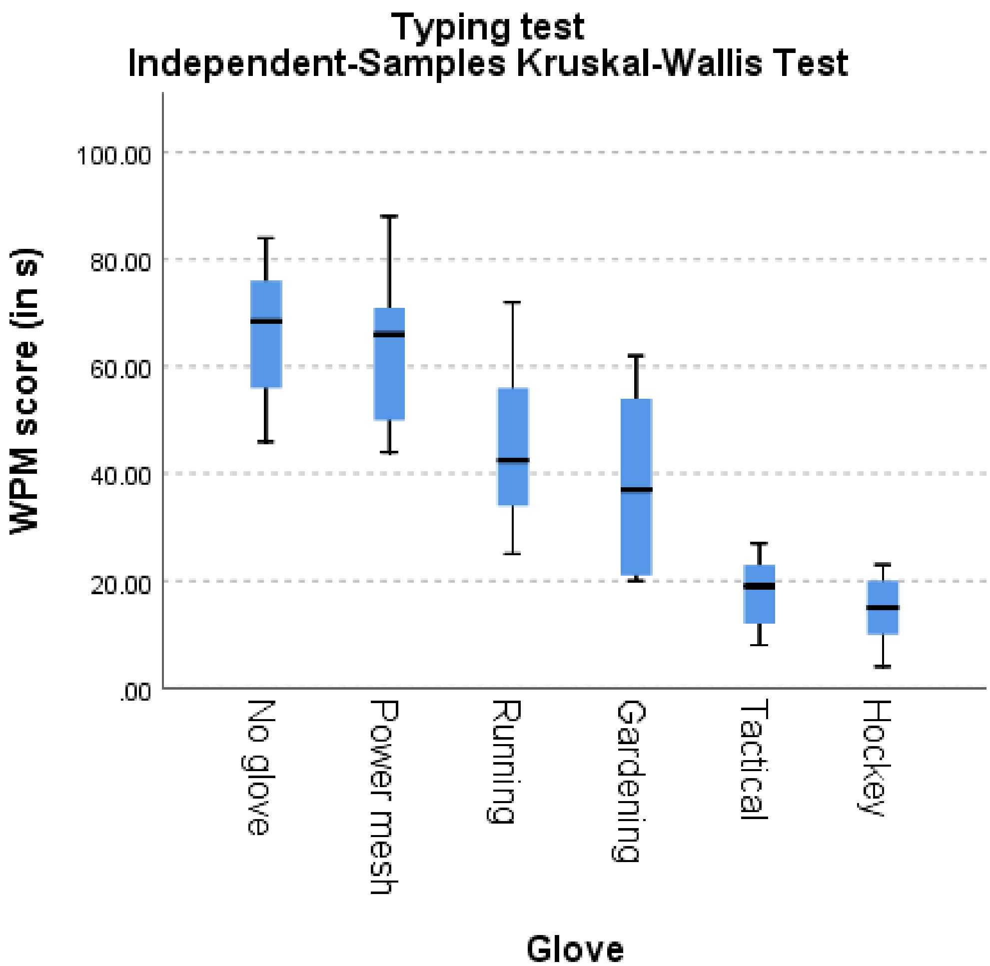


Figure 53 – Typing test Kruskal-Wallis test results – there is a significant difference in the WPM score between gloves, $\chi 2(5) = 26.05$, $p = 0.000$

The typing test in Study 2 largely replicated the results from Study 1, in that there was a consistent decrease in the WPM performance as the gloves increased in their bulk. The results from Study 2 – seen in Figure 53 – showed a median score of ~70 WPM with the "no glove" condition, falling to a median of ~15 WPM with the Hockey gloves. This is a notable result as the participant population in Study 1 had a median “no glove” score of ~50 WPM, which fell to ~10 WPM with the Hockey glove. While the Study 2 population had a noticeably different baseline score, both tests showed similar decreases in their typing skills as the glove bulk increased. The Tactical and Hockey gloves again showed comparatively little variability in their WPM scores compared to the other gloves.

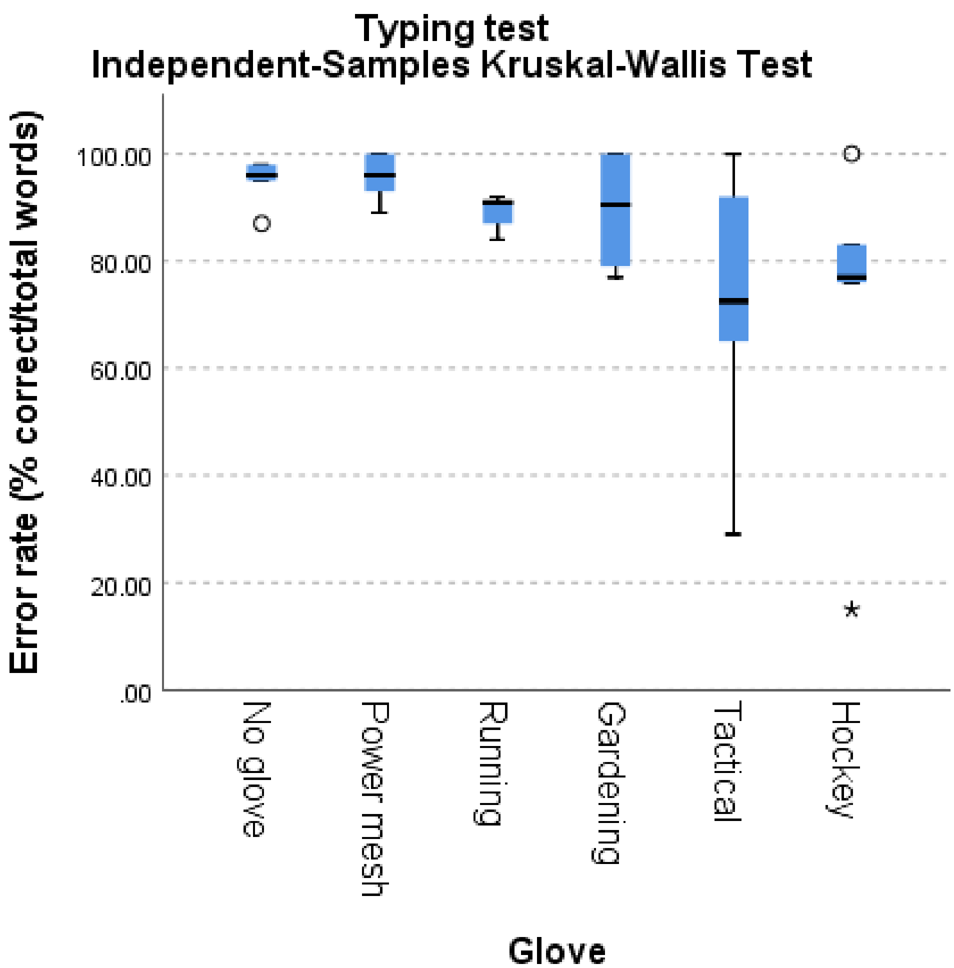


Figure 54 – Typing test error rate Kruskal-Wallis test results – there is a significant difference in the error rate between gloves, $\chi 2(5) = 11.39$, $p = 0.044$

However, as seen in Figure 54, the Study 2 typing test showed increased variability in the error rate of the Tactical and Hockey gloves, with some extreme outliers. As the results range from 100% of the words being correctly typed to less than 20% correct, it is clear that the gloves had a significant effect on the participant's ability to maintain their normal typing skill.

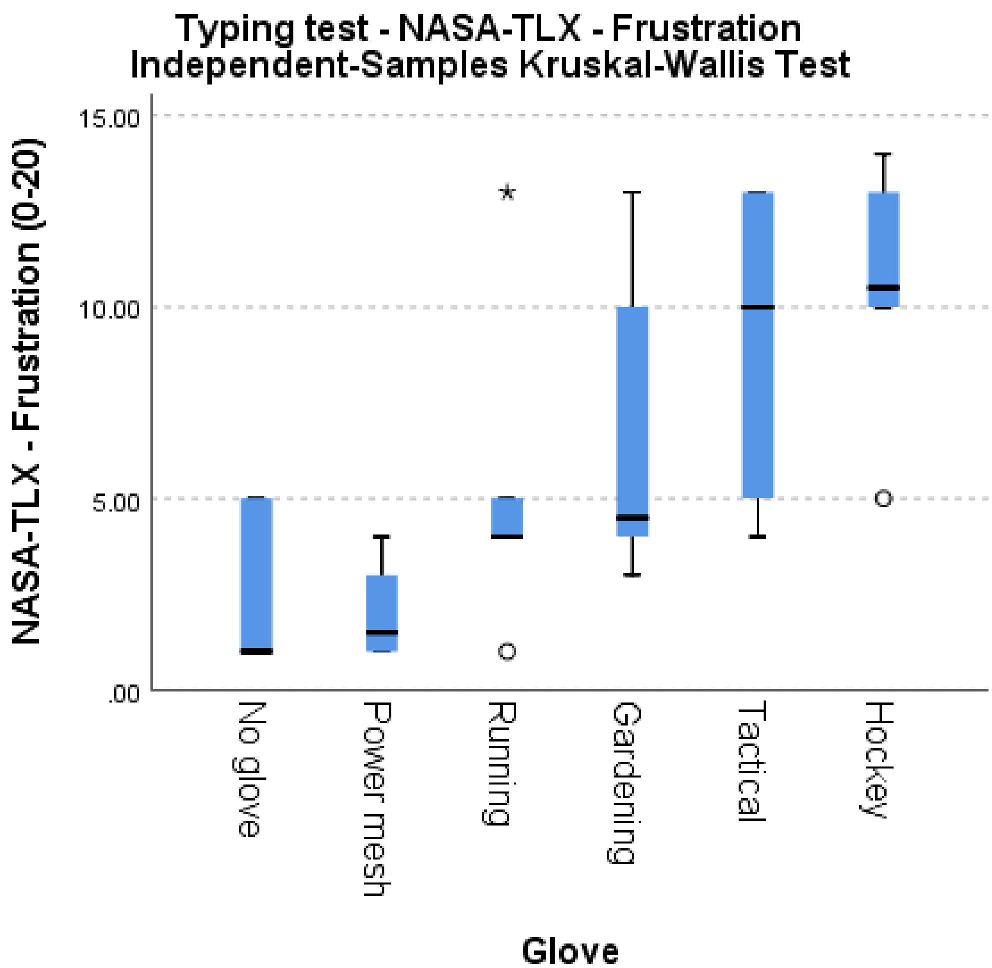


Figure 55 – Typing test NASA-TLX Frustration scale Kruskal-Wallis test results – there is a significant difference in the Frustration score between gloves, $\chi 2(5) = 19.61$, $p = 0.001$

There is also a corresponding change in the perceived frustration of completing this task, seen in Figure 55. Participants reported that they could not feel the keyboard while wearing the gloves, and that they could not tell they were positioned over the keys until they had already started to push. They also reported feeling like they had lost control over the task while wearing gloves, that they needed to constantly visually check the status of the keys, and that they needed to adopt a hunt and peck strategy with the Hockey gloves. It is clear that these disruptions led to their increased frustration with the task.

6.3.2 Bennett hand tool test

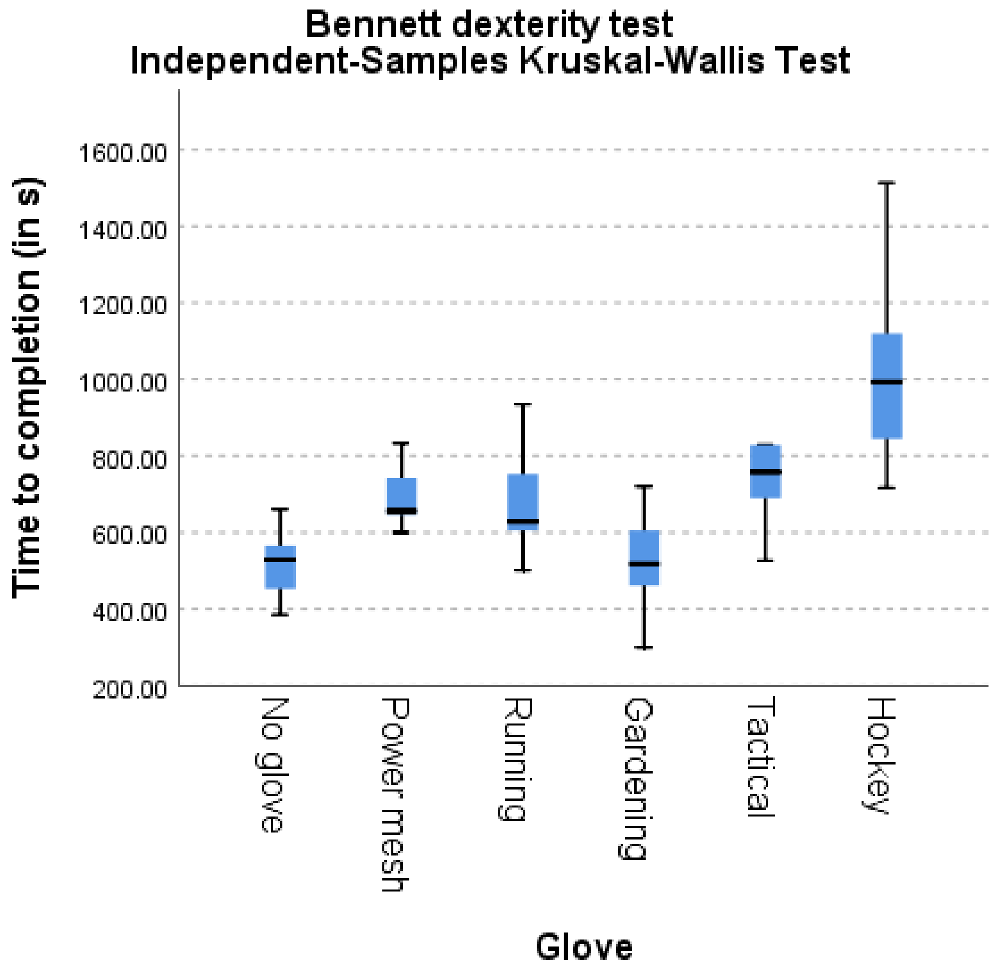


Figure 56 – Bennett dexterity test Kruskal-Wallis test results – there is a significant difference in the time to completion score between gloves, $\chi2(5) = 20.2$, $p = 0.001$

The Bennet dexterity test showed significant differences between the gloves, shown in Figure 56. The “No glove” condition had a median time score of ~525 seconds to complete the task, while the Hockey glove nearly doubled that result with a median score of ~1000 seconds. Notably the Gardening glove did very well, roughly equaling the “No glove” score, with some participants scoring better with the Gardening glove.

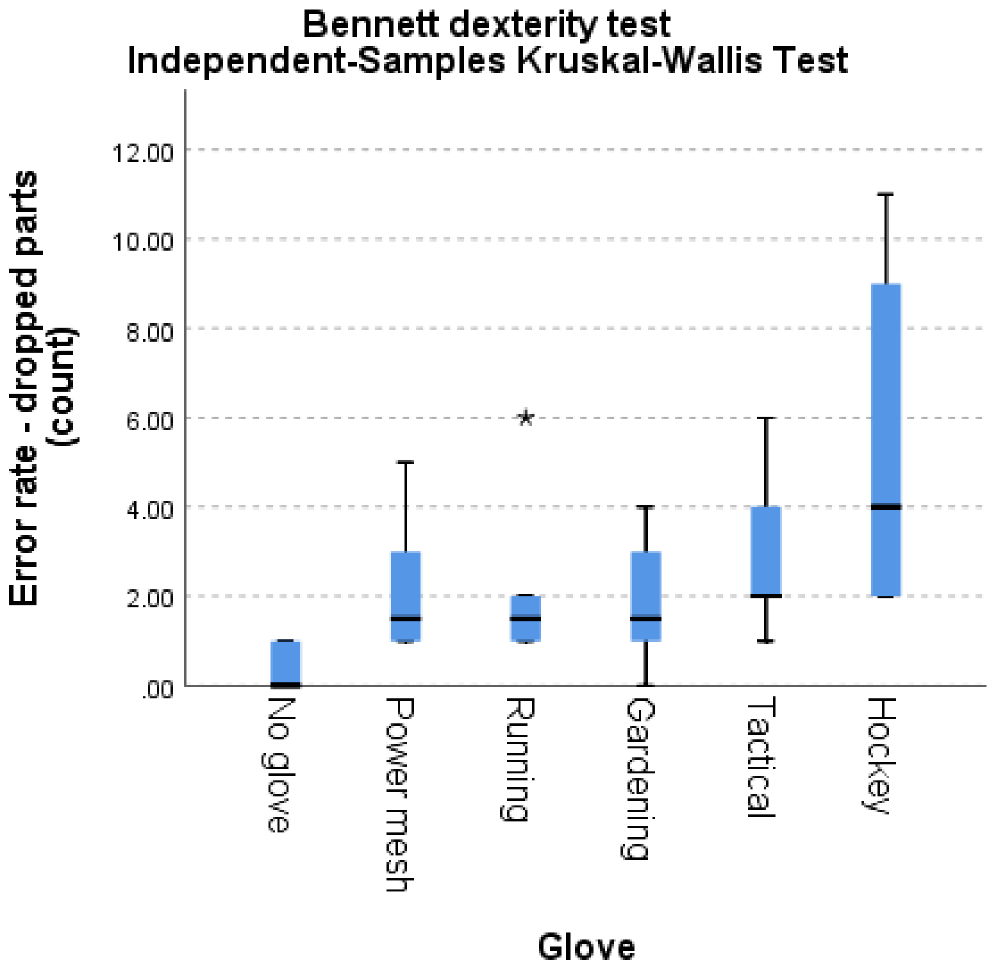


Figure 57 – Bennett dexterity test error rate Kruskal-Wallis test results – there is a significant difference in the error rate between gloves, $\chi2(5) = 16.30$, $p = 0.006$

The Bennet test again showed significant differences in the number of parts dropped while wearing each glove. As shown in Figure 57, participants dropped parts with the “no glove” condition with a median rate close to zero, while they dropped parts at a median rate of 4 times during the session with the Hockey gloves. The variability of the Hockey glove error score was also much higher than other conditions, with some participants dropping up to 11 parts during the task.

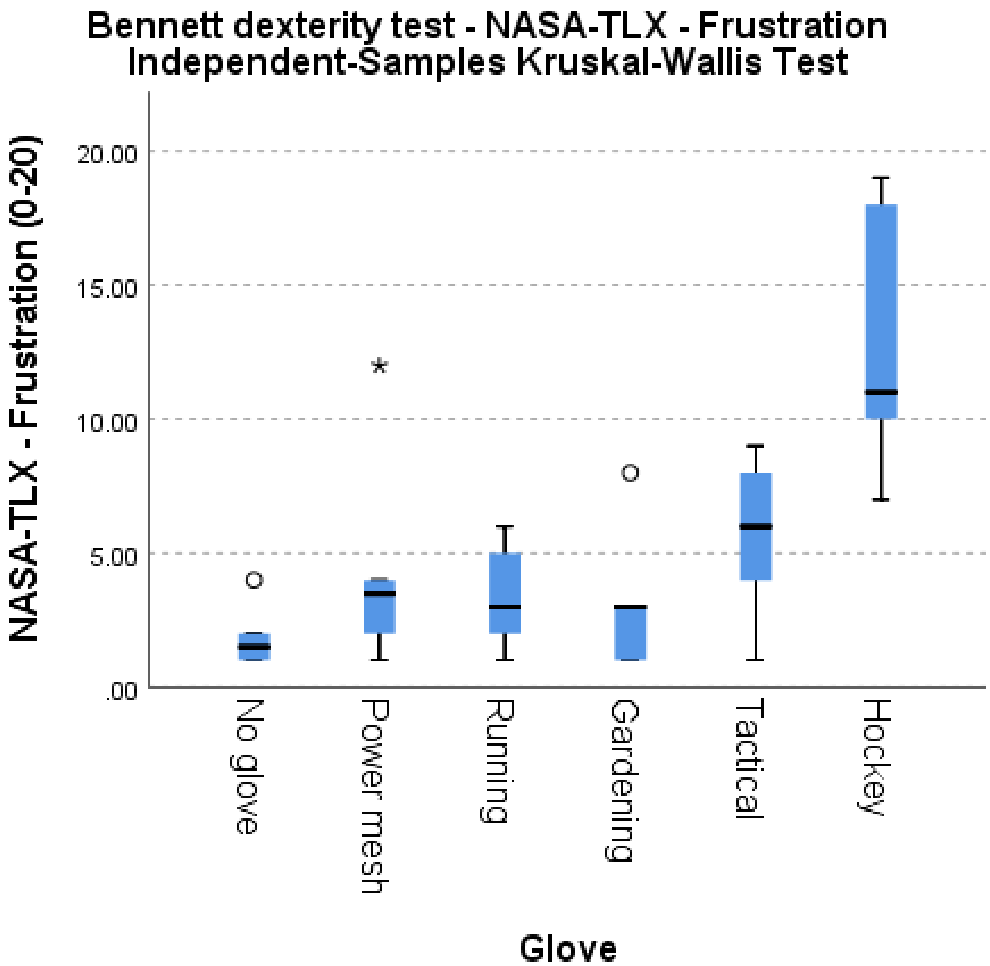


Figure 58 – Bennett dexterity test NASA-TLX Frustration scale Kruskal-Wallis test results – there is a significant difference in the Frustration scale between gloves, $\chi2(5) = 17.56$, $p = 0.004$

This variability, seen again in Figure 58, showed in the NASA-TLX Frustration score as well. The Hockey gloves contributed to a significant increase in the perceived frustration with the task. Participant's attributed this to the bulk of the gloves getting in their way during fine motor tasks, blocking their view of the parts due to the visual occlusion of the gloves, and the need to move their entire hand to manipulate smaller objects. By contrast, the Gardening glove was praised for its enhanced grip, which allowed for consistent control of the small parts.

6.3.3 O'Conner Test

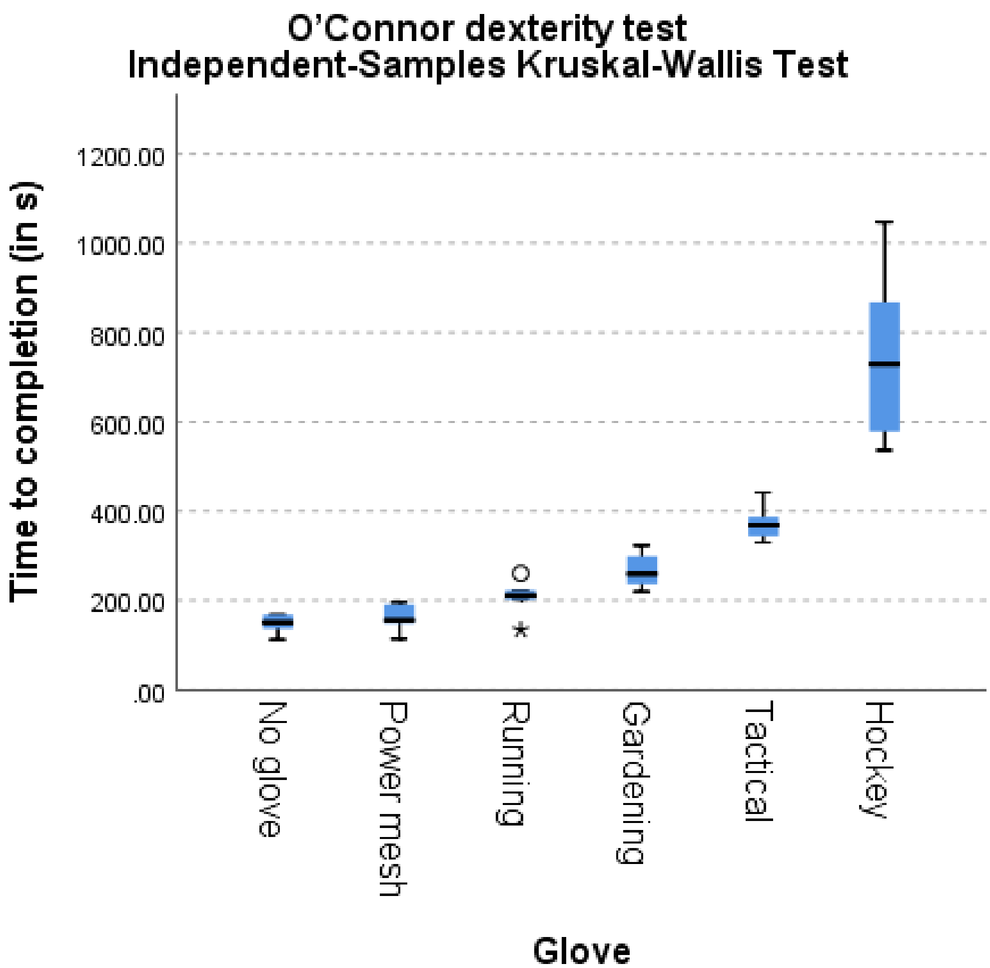


Figure 59 – O'Conner dexterity test Kruskal-Wallis test results – there is a significant difference in the time to completion score between gloves, $\chi2(5) = 31.26$, $p = 0.000$

Results from the O'Conner test, seen in Figure 59, showed a significant difference in performance between the gloves, with a consistent curve of the median scores. The "no glove" condition was the fastest with a median score of ~174 seconds to complete the task, while the Hockey showed a median score of over 700 seconds, with a high range of variability.

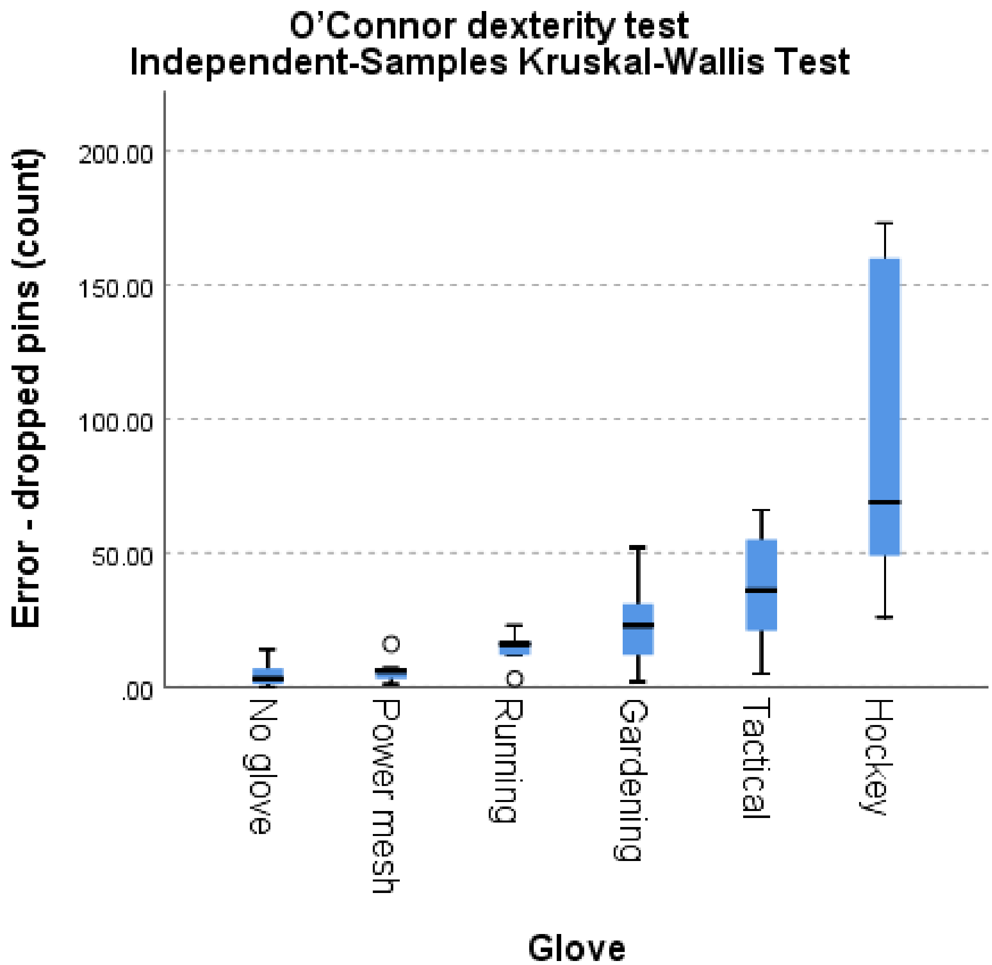


Figure 60 – O'Conner dexterity test error rate Kruskal-Wallis test results – there is a significant difference in the error rate between gloves, χ2(5) = 22.30, p = 0.000

The difference in error rate is also significant, as shown in Figure 60. The "no glove" condition showed the participants dropping pins during the task at a median rate close to zero, while the Hockey gloves resulted in a median rate of ~65 pins, while some participants dropped more than 150 on their way to completion of the task. Given that the completing the task required placing 150 total pins, this was an error rate in excess of 100%, compared to the correctly placed pins, and shows the difficulty in completing the tasks with bulky gloves.

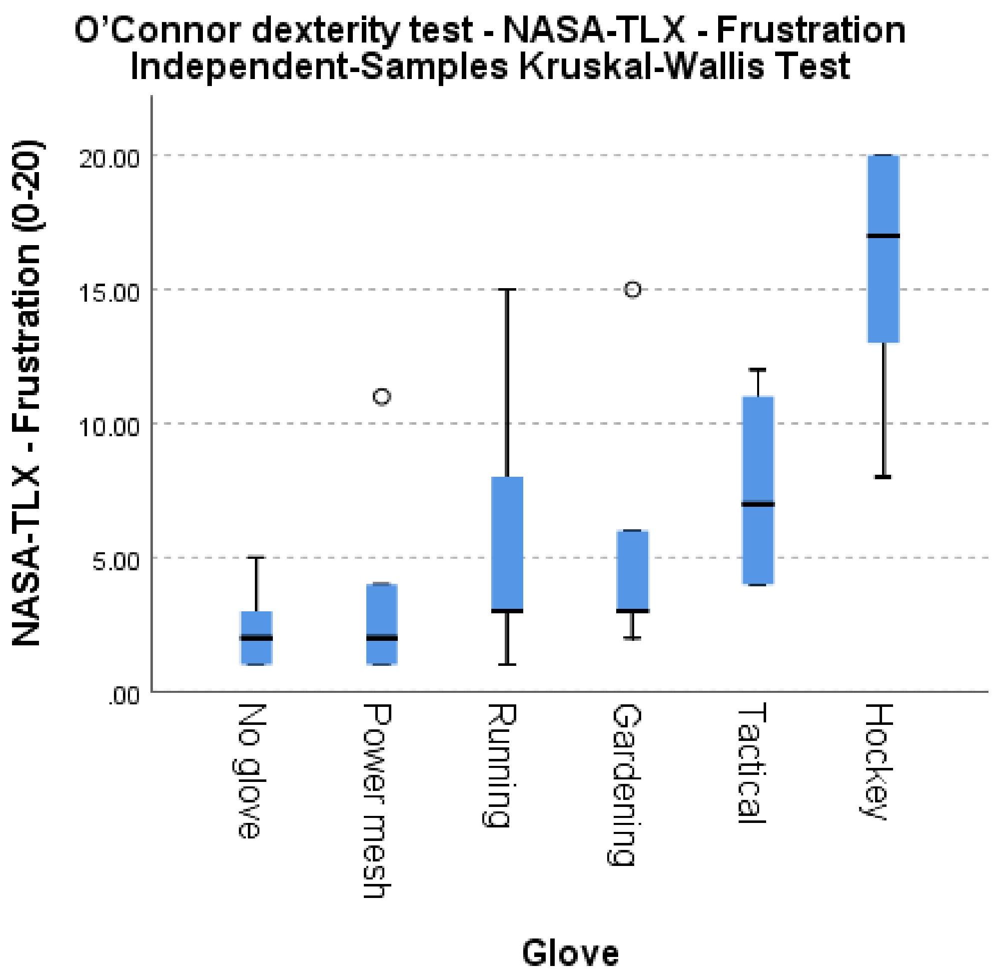


Figure 61 – O'Connor dexterity test NASA-TLX Frustration scale Kruskal-Wallis test results – there is a significant difference in the Frustration scale between gloves, $\chi 2(5) = 18.138$, $p = 0.003$

These results are again reflected in the perceived Frustration of the task – with Figure 61 showing the Hockey gloves with a very high rate of frustration, compared to most of the other gloves. Participant's attributed this frustration to their inability to see the pins in their hands, to control the number of pins they were picking up, and to successfully grip the small pins between their thumb and index finger – many reported switching to using their index and middle finger instead. Participants also reported that many found the pins sticking into the fabric of the glove and found it very hard to remove them. The notable frustration score for the Running Gloves – which performed well on the other metrics – can be attributed to participant's dislike for the conductive fabric swatches on the thumb and index fingers, which were intended to ease operation of capacitive screens on mobile phones. In this case participants found them to be slippery and found that they further inhibited their ability to feel the count of the pins in their fingers.

6.3.4 Minnesota test

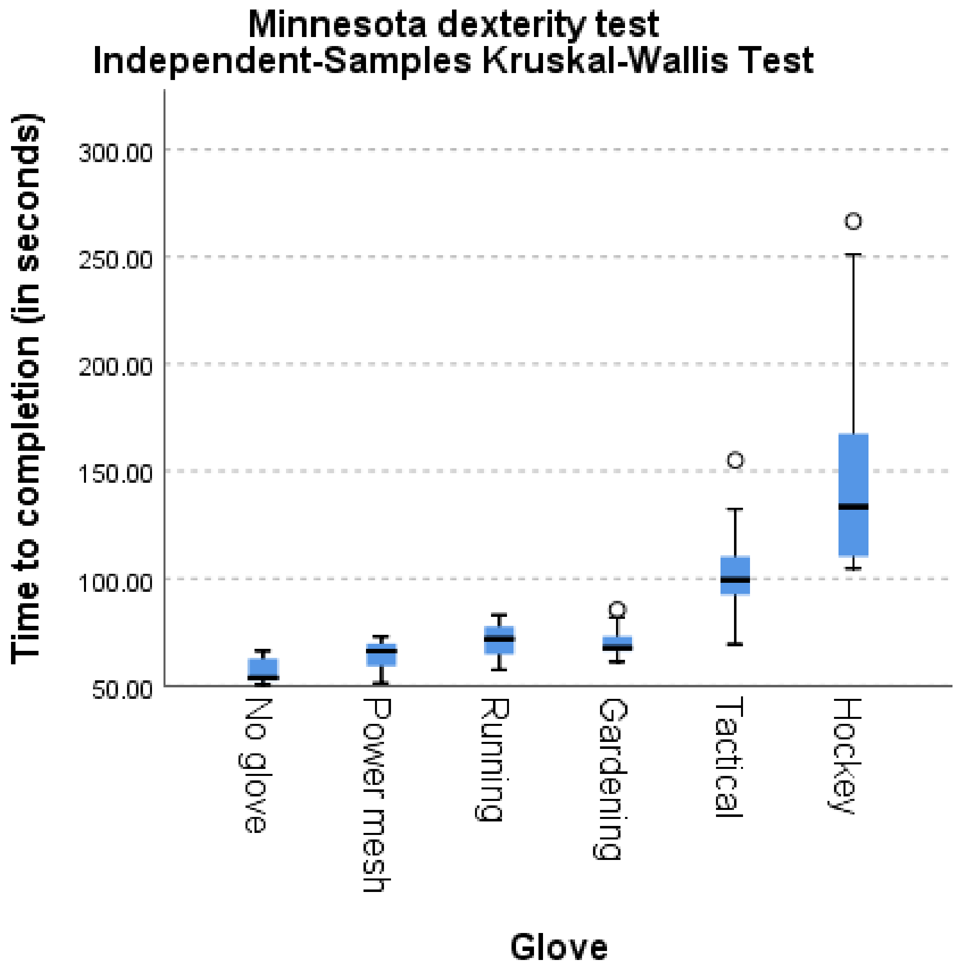


Figure 62 – Minnesota dexterity test Kruskal-Wallis test results – there is a significant difference in the time to completion score between gloves, $\chi2(5) = 54.374$, p = 0.000

The Minnesota dexterity test showed a significant difference in the time to completion between gloves, seen in Figure 62. The “no glove” condition allowed participants to complete the task of turning the 60 disks at a median time just above 50 seconds, while the Hockey gloves required a median time of ~130 seconds to complete the same task. The

Hockey glove again showed increased variability, as some participants required over 250 seconds to complete the task.

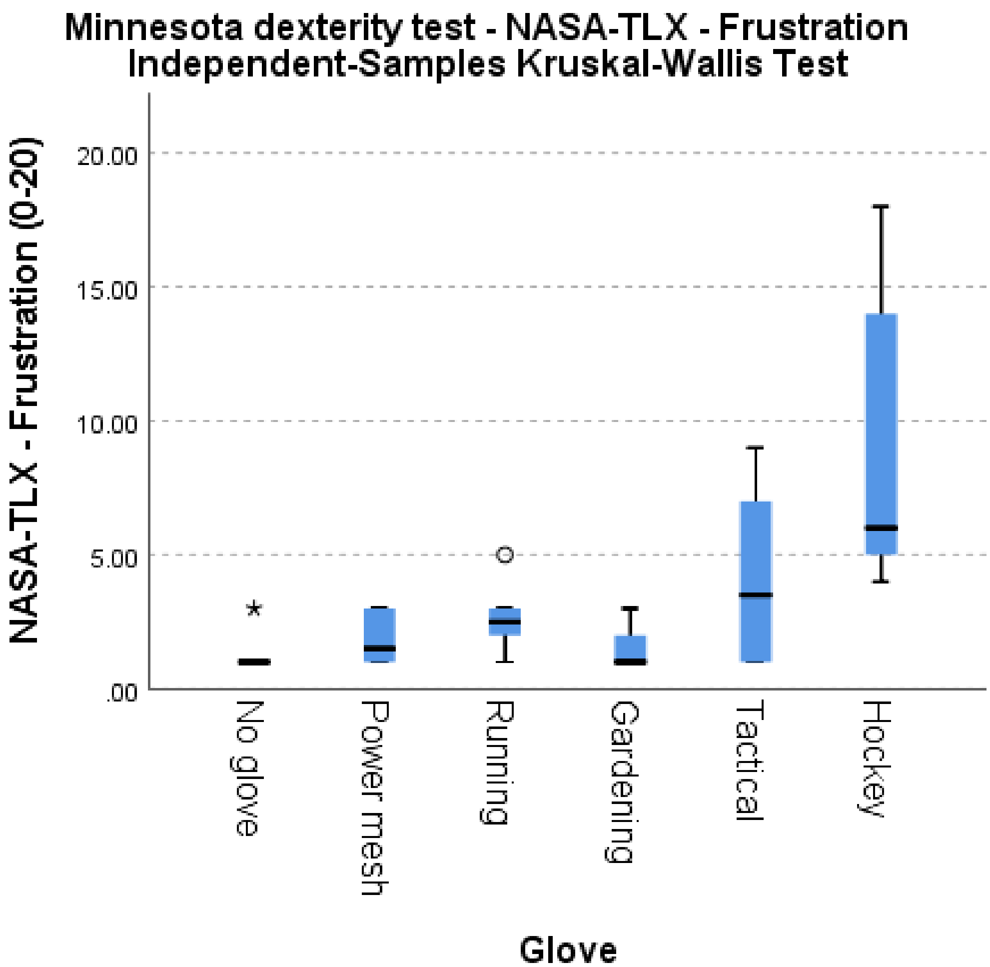


Figure 63 – Minnesota dexterity test NASA-TLX Frustration scale Kruskal-Wallis test results – there is a significant difference in the Frustration scale between gloves, $\chi2(5) = 18.447$, $p = 0.002$

The perceived frustration of this task showed results seen in Figure 63 that mirror the time results, with the Hockey glove scoring significantly higher on the NASA-TLX scale. Participants attributed this to the Hockey gloves occluding their view of the disks, limiting the ability to pick up the disks, and forcing them to develop a new strategy for the task.

6.3.5 Tactile Discrimination test

Table 6 – Hypothesis test summary for Kruskal-Wallis test across all six Tactile discrimination test conditions. Significance level is 0.05.

Task Condition	Test Statistic	Sig.
(A) Circle	12.35	.030
(B) Oval	6.04	.302
(C) Hollow Square	15.24	.009
(D) Star	6.29	.279
(E) Hexagon	12.71	.026
(F) Square	11.17	.048

The Tactile Discrimination test showed significant differences between the gloves in four of the six task conditions, with the Circle, Hollow Square, Hexagon, and Square all showing significant differences (Table 6). These differences are in the scale of the shapes participants were able to correctly identify for each task condition.

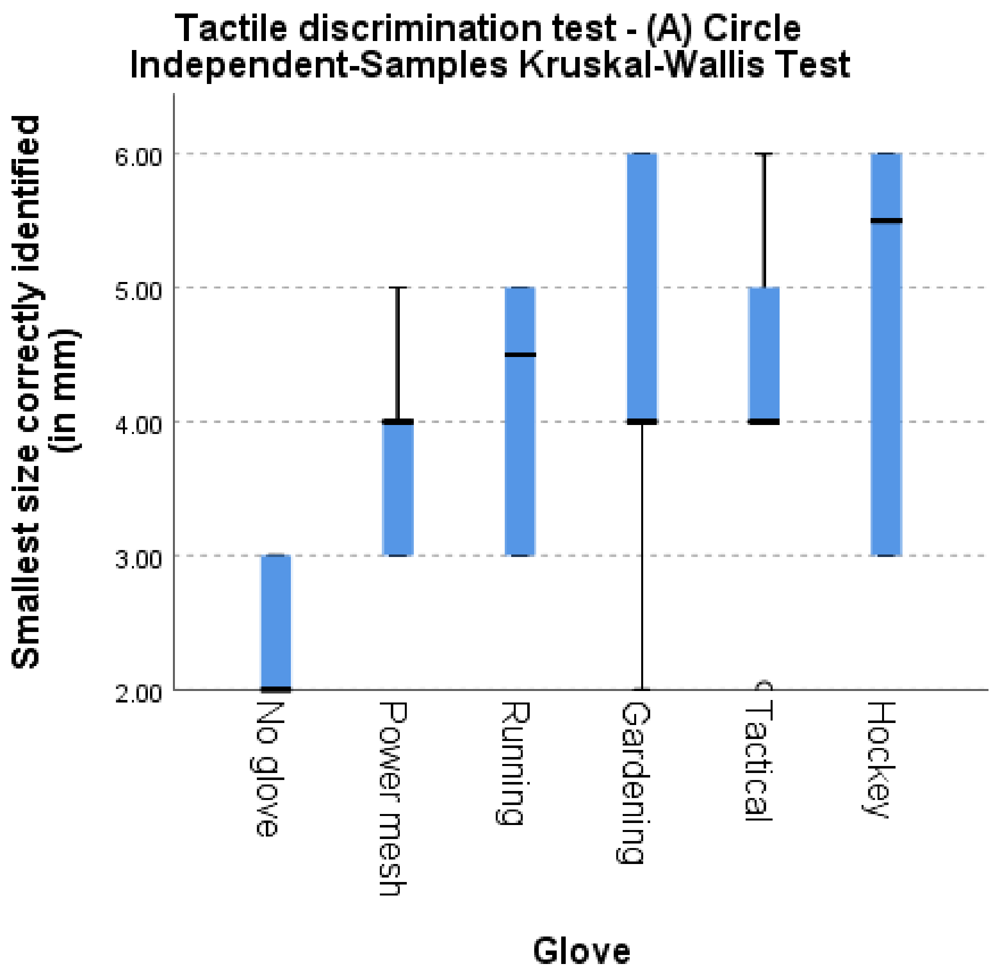


Figure 64 – Tactile discrimination test Kruskal-Wallis test results – there is a significant difference in the scale of the smallest size of the Circle shape correctly identified between gloves, $\chi2(5) = 12.35$, $p = 0.030$

Using the Circle condition as an example, shown in Figure 64, participants were able to reliably identify the smallest scale in the test (2mm) with the "no glove" condition, while the median result with the Hockey glove was between the 5mm and 6mm size. All of the gloves tested showed a median size of 4mm or above for this test, showing the relative impairment that a relatively thin piece of fabric can have on small-scale sensing tasks. The Gardening gloves thicker Nitrile dip coating hindered some participants, as they reported the rubber coating disrupted their ability to feel edges, which cause them to confuse the circle and other shapes more easily.

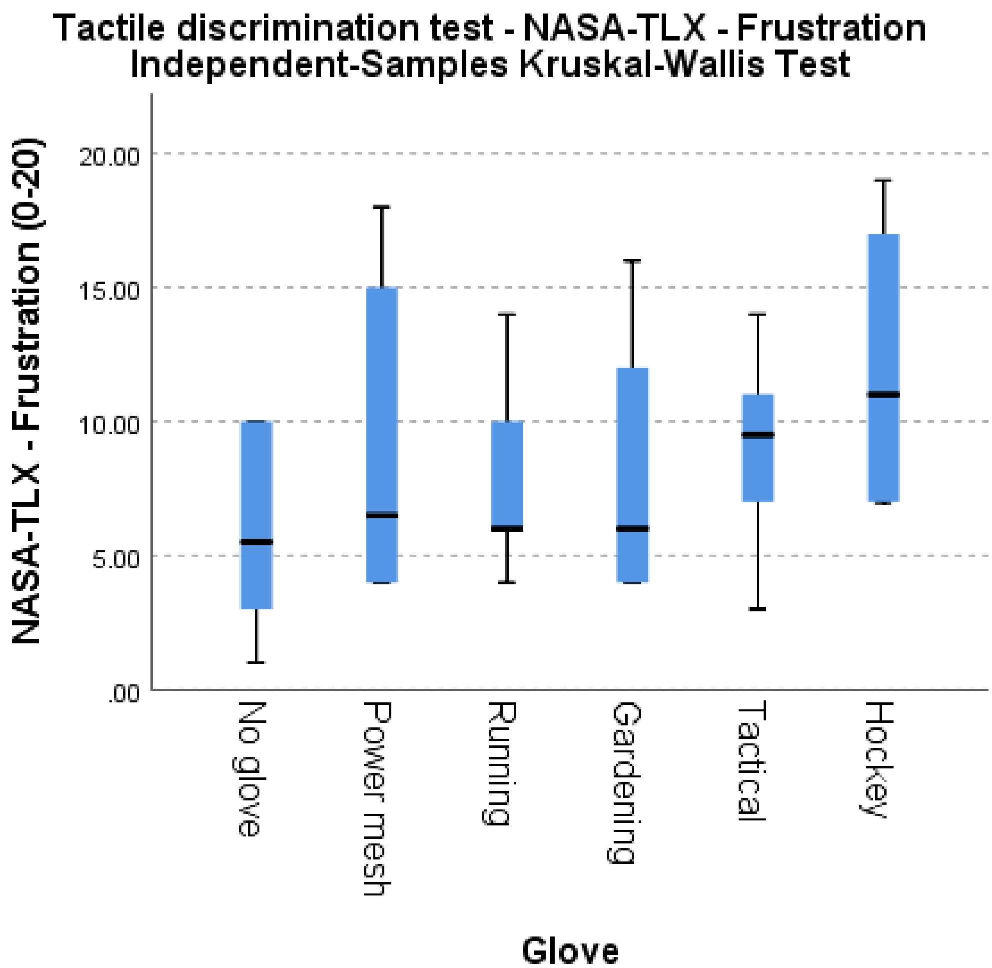


Figure 65 – Tactile discrimination test NASA-TLX Frustration scale Kruskal-Wallis test results – there is no significant difference in the Frustration scale between gloves, $\chi 2(5) = 5.71$, $p = 0.335$

The NASA-TLX Frustration scale did not show a significant result between gloves for this task. It is notable that the participant's performance scores showed differences, but they did not perceive those differences as contributing to increased frustration during the task.

6.4 Glove findings

6.4.1 Scaled response – Typing questions

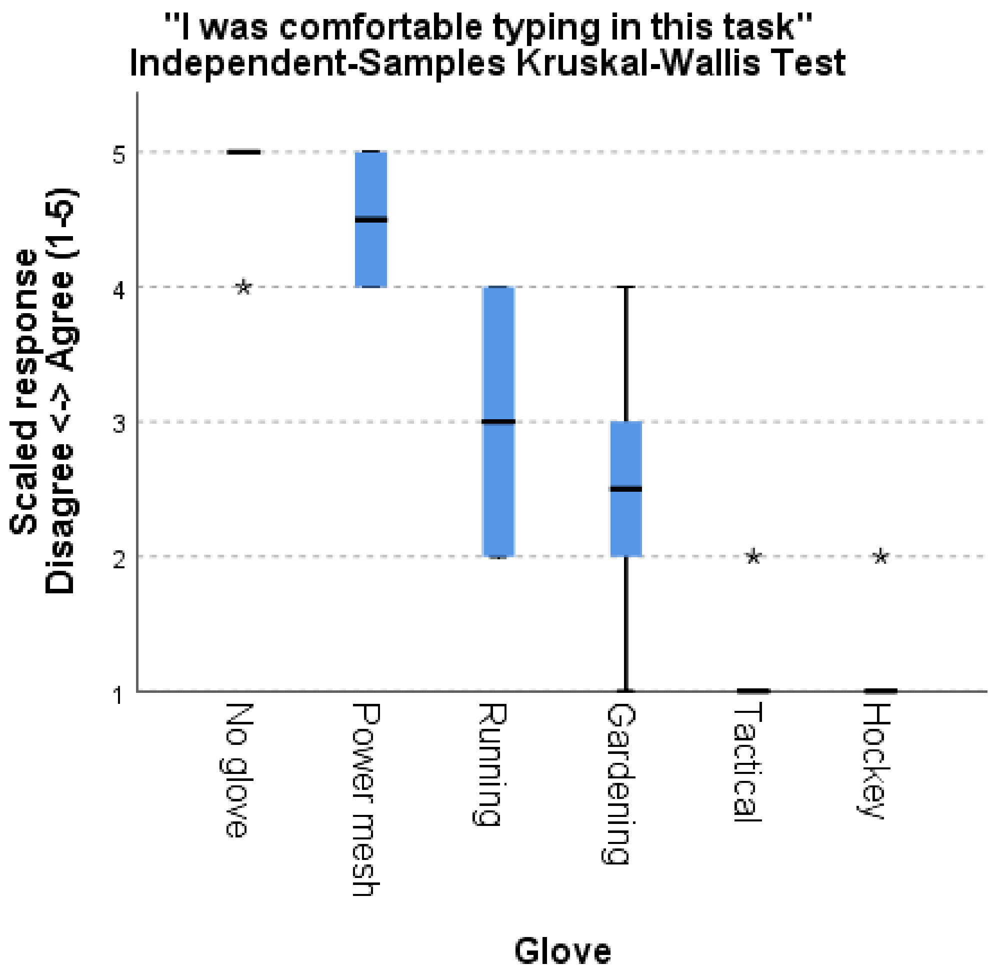


Figure 66 – Scaled response for "I was comfortable typing in this task" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi2(5) = 29.11$, $p = 0.000$

The typing test questions in Study 2 replicate the results in Study 1, showing significant differences between gloves they affected the participant's perceived comfort, normal typing ability, and eye position. For comfort, the Powermesh gloves again were strongly preferred, as the Tactical and Hockey gloves received low marks – seen in Figure 66. Participant's explanation for this were the Tactical gloves restrictive materials, and the Hockey gloves forcing them into awkward hand positions.

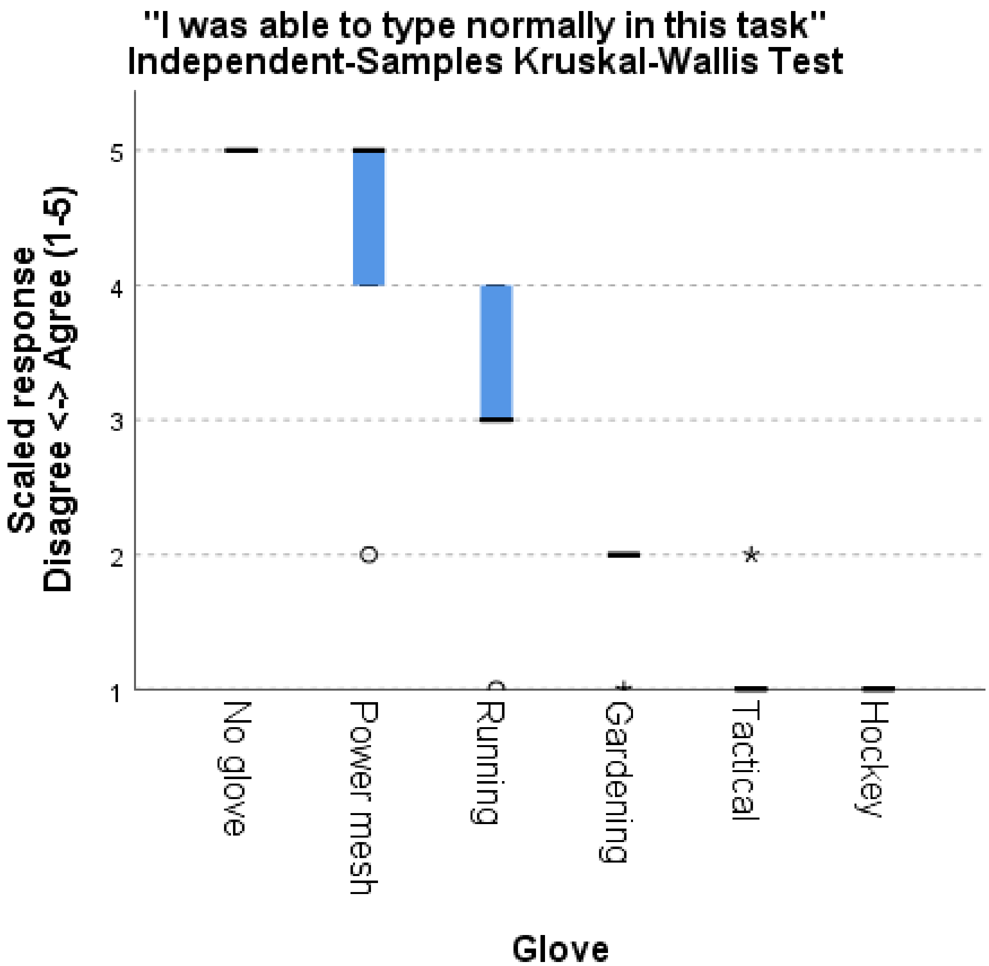


Figure 67 – Scaled response for "I was able to type normally in this task" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi 2(5) = 29.39$, p = 0.000

Participant's normal typing capabilities were again adversely affected by the Tactical and Hockey gloves, shown in Figure 67. For the Hockey glove, the overall bulk of the gloves, and the long fingers that kept participant's hands away from the keys were cited as common experiences. Participants felt like they needed to dramatically slow down and type very gently to avoid hitting multiple keys at the same time, and most reverted to a hunt and peck methods when their touch typing was interrupted.

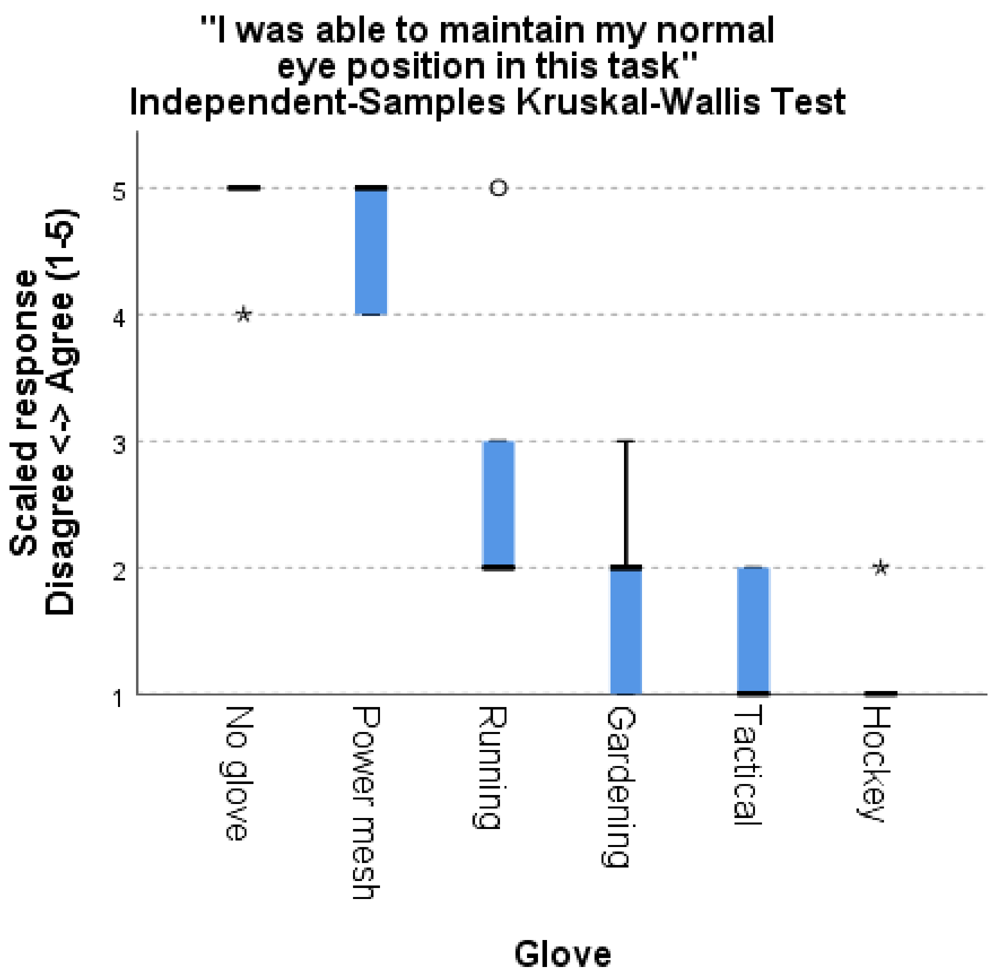


Figure 68 – Scaled response for "I was able to maintain my normal eye position in this task" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi2(5) = 27.85$, $p = 0.000$

This disruption again showed some effect on their ability to maintain a normal typing gaze position. As shown in Figure 68 the Hockey gloves required participants to focus their attention on the keyboard to ensure they were hitting the right key but needing to switch their gaze to the screen to confirm that they had been successful. Participants reported they this strategy required them to memorize the word from the screen, attempt to pick out the letters with their gaze on the keyboard, and then revert their gaze back to the screen to confirm the results. This caused extra mental load for them, as they would not normally be trying to memorize the words as they typed.

6.4.2 Scaled response – Comfort and fit

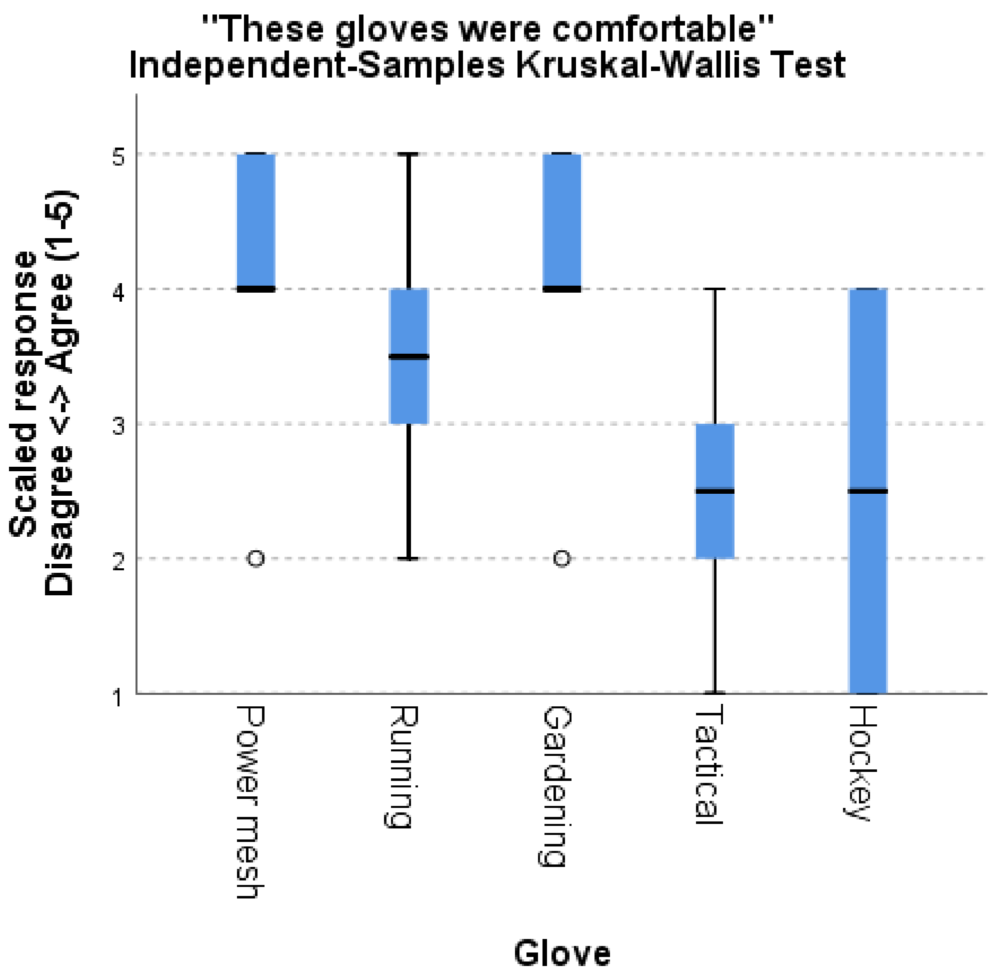


Figure 69 – Scaled response for "These gloves were comfortable" Kruskal-Wallis test results – there is no significant difference in agreement with the statement between gloves, $\chi 2(4) = 8.89$, p = 0.064

Participants in Study 2 did not report significant differences between gloves related to their perceived comfort, as they felt most of the gloves were of middling comfort. The Powermesh was again favoured, as shown in Figure 69, though this group also thought the Gardening gloves were more comfortable than the rest of the set.

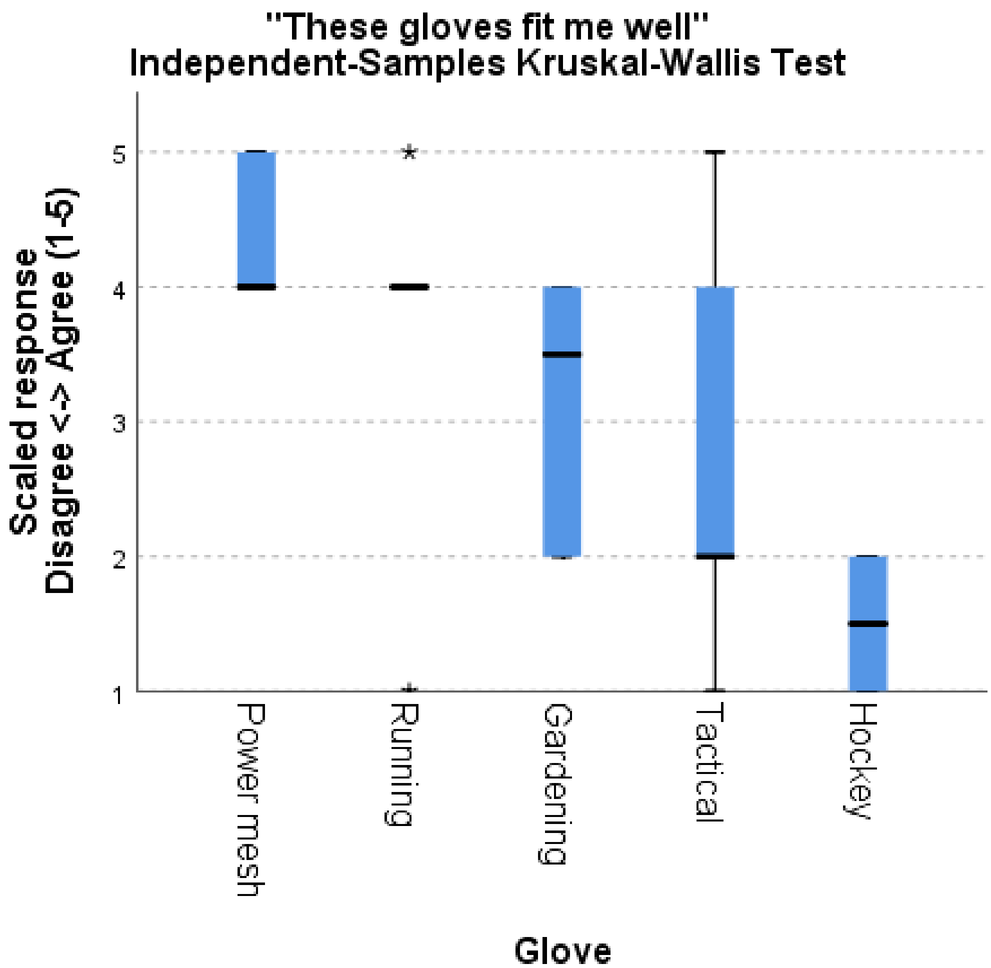


Figure 70 – Scaled response for "These gloves fit me well" Kruskal-Wallis test results – there is no significant difference in agreement with the statement between gloves, χ2(4) = 13.63, p = 0.009

There were significant results in the perceived differences between glove fit, however. Figure 70 shows the Powermesh gloves were preferred for their conforming and extensible fit, as the mesh adjusted easily to fit most of the participant's hands. They singled out the Tactical gloves with the opposite effect, as the stiff material in these gloves made the fit far less forgiving. The Hockey gloves again fit almost no one in the task, and their long fingers kept participant's hands away from the objects they were interacting with.

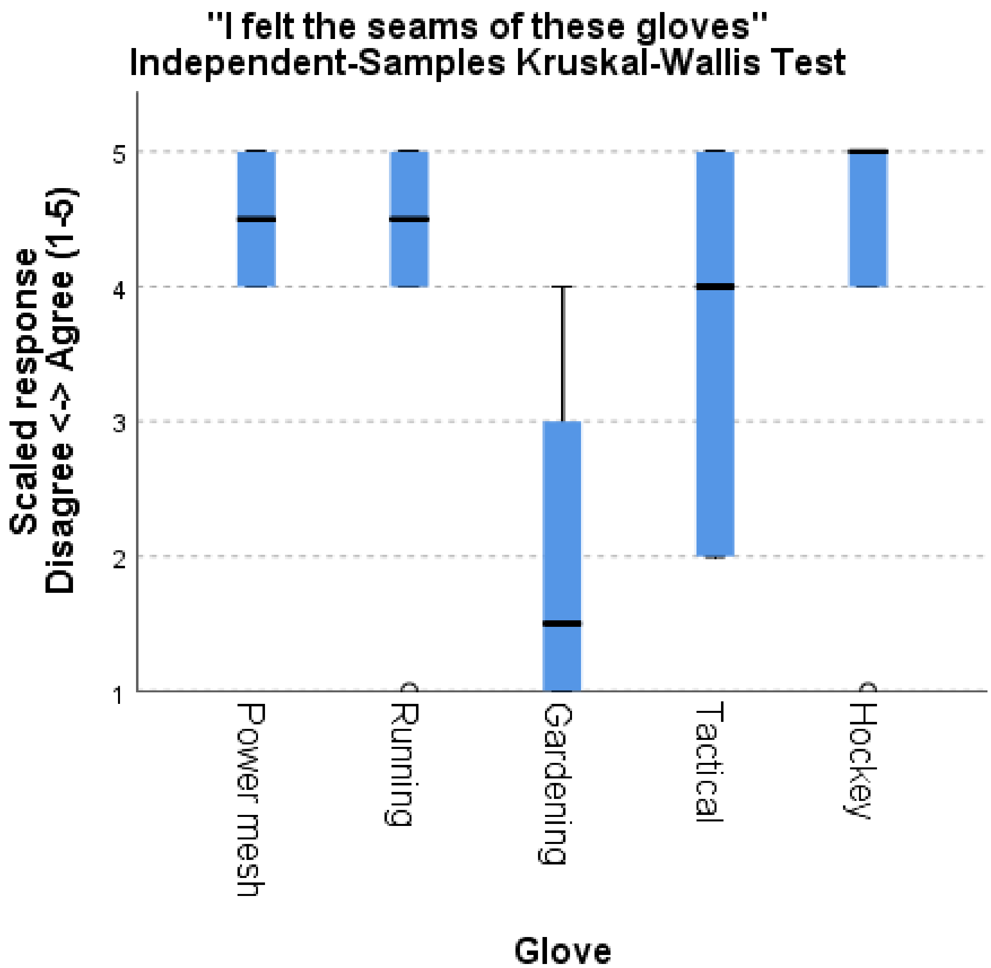


Figure 71 – Scaled response for "I felt the seams of these gloves" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi 2(4) = 9.61$, $p = 0.047$

This question around seams again singles out the Gardening gloves with a significant result, shown in Figure 71. As the Gardening gloves are the only ones in the set built without seams in their construction, this is an accurate assessment from the participants, and shows how the small details of the seams can be perceived.

6.4.3 Scaled response – Slip and grip

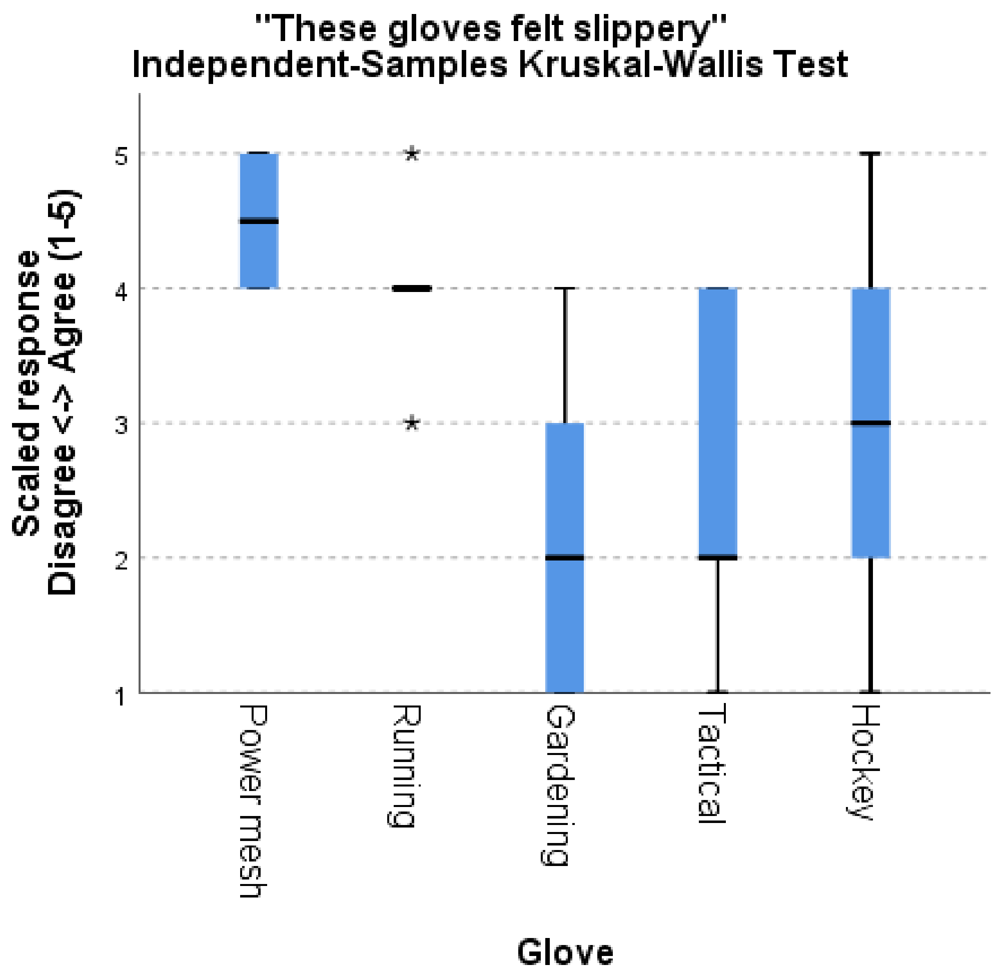


Figure 72 – Scaled response for "These gloves felt slippery" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi2(4) = 13.54$, p = 0.009

The slip/grip questions again replicate the results from Study 1, with the three groups emerging. The slippery question shows a significant difference between gloves in Figure

72, with the Gardening gloves seen as the least slippery. The Tactical and Hockey gloves follow as the second group of middling slipperiness, with the Running and Powermesh gloves seen as the most slippery.

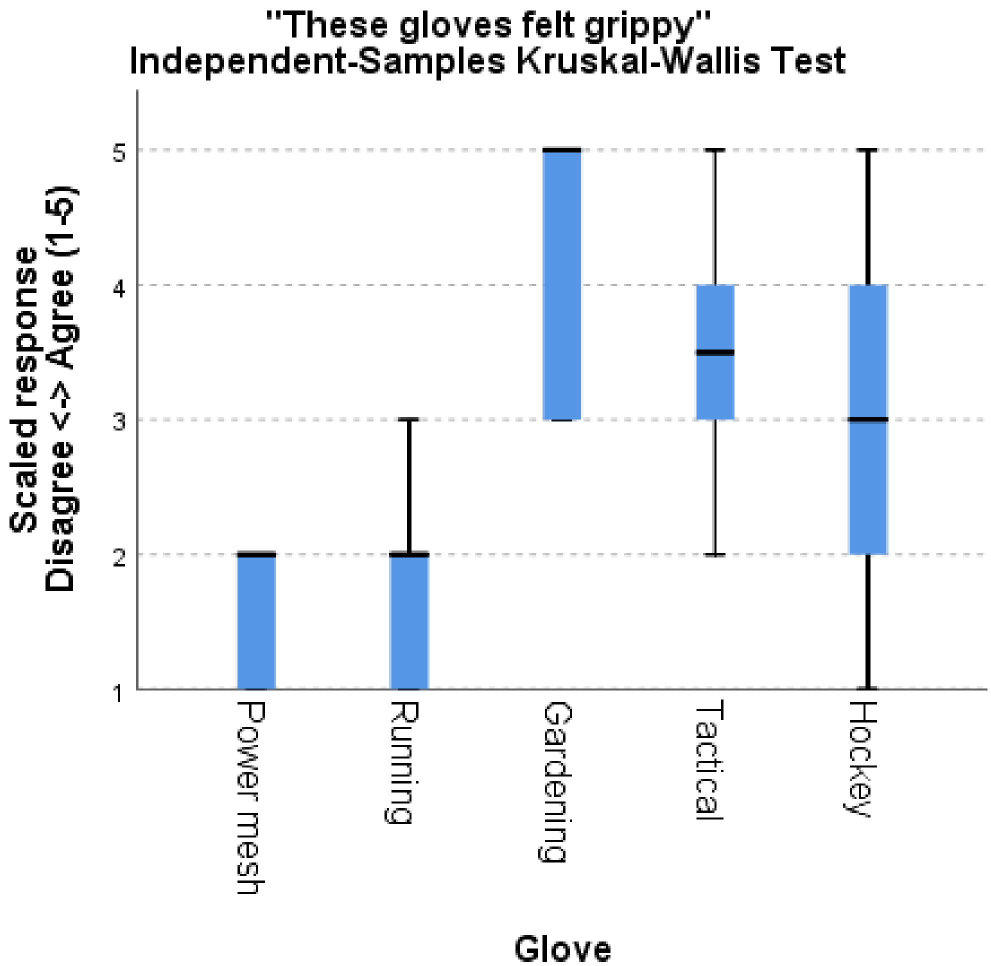


Figure 73 – Scaled response for "These gloves felt grippy" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi2(4) = 15.78$, $p = 0.003$

The grip questions show significant results which reverse this order in Figure 73. The Gardening glove's nitrile dip coating is the only surface finish tested that reliably provides extra grip to make up for the loss of grip on the wearer's skin.

6.4.4 Scaled response – Hand mobility

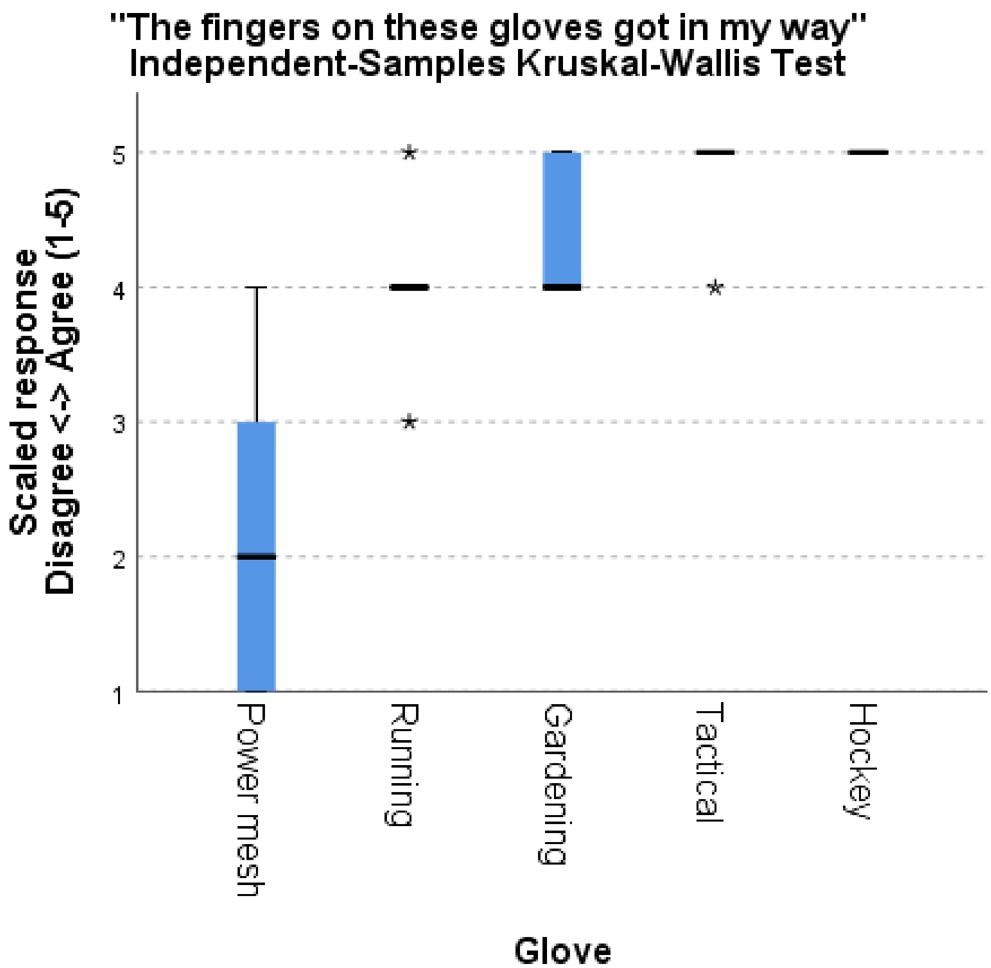


Figure 74 – Scaled response for "The fingers on these gloves got in my way" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi2(4) = 20.67$, $p = 0.000$

The final set of questions is focused on hand mobility and physical encumbrance. The first result shows in Figure 74 a significant difference between gloves in their ability to get in the participant's way during a task. The less-restrictive Powermesh gloves score well here, with the bulky and ill-fitting Tactical and Hockey gloves scoring poorly.

Figure 75 – Scaled response for "I had difficulty moving my hands in these gloves" Kruskal-Wallis test results – there is no significant difference in agreement with the statement between gloves, $\chi2(4) = 7.34$, $p = 0.119$

Study 2 did not show a significant difference between gloves in limiting the participant's hand movements. While the Powermesh gloves and Hockey gloves scored quite differently, there was a high degree of variability in the participant's responses across the gloves, which led to an insignificant result.

6.5 Discussion

6.5.1 Glove preference

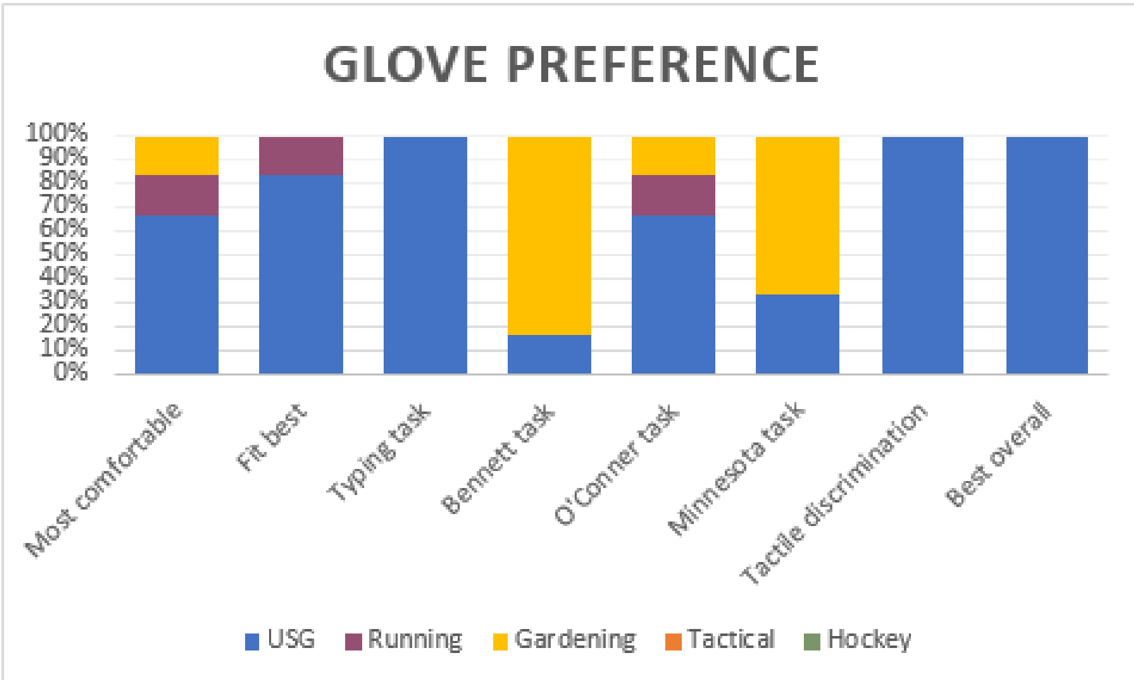


Figure 76 – Aggregated answers for post-session interview questions about preferred glove choice for comfort, fit, and task completion.

At the conclusion of the study, participants were asked to indicate their preferred glove for fit, comfort, task completion for each task, and their overall choice to complete the entire study with. For Study 2, the USG/Powermesh glove dominated user preference across all categories with the exception of two, as shown in Figure 76. The Bennet dexterity test and the Minnesota dexterity task both have an emphasis on dexterous interaction with objects, and less of an emphasis on sensing. In these cases, participants seemed to prefer the added grip of the Gardening gloves, as they were given a more secure and reliable grip on the objects. In all other cases the stretchy fit and sensory transparency of the USG/Powermesh glove were seen to provide superior capabilities to the wearer across most tasks, which led to the unanimous choice of this glove as the overall winner.

Table 7 – Aggregated values for post-session interview questions, with totals. Values shown indicate the number of participants who picked each glove as their preferred option for each question about comfort, fit, and task completion.

Preference	USG	Running	Gardening	Tactical	Hockey
Most comfortable	4	1	1	0	0
Fit best	5	1	0	0	0
Typing task	6	0	0	0	0
Bennett task	1	0	5	0	0
O'Conner task	4	1	1	0	0
Minnesota task	2	0	4	0	0
Tactile discrimination	6	0	0	0	0
Best overall	6	0	0	0	0
Total	34	3	11	0	0

Looking at the numerical breakdown of the selections (Table 7), it is clear that users would benefit from the USG/Powermesh glove even more if it had the enhanced grip capabilities of the Gardening glove. This is a consistent finding from Study 1 and will inform the glove selection for Study 3. It is also clear that none of the attributes of the Tactical and Hockey gloves were preferred by participants, as they received no votes for these questions.

6.5.2 Task selection

Study 2 was intended to update the protocol used by Study 1 with the intent of replacing tasks that had been developed in-house with tasks that were well reviewed and accepted by the broader community. This led to the introduction of the Bennet dexterity test, O’Conner dexterity test, and Minnesota dexterity test, all of which were trialled for the first time in this study. It seems clear that these new tests performed well and were able to show similarly significant differences within the Commercial Glove set to the prior protocol, while providing a higher degree of external validity and acceptance.

Each test provided an important improvement on its predecessor. The Bennet dexterity test provided a similar dexterity task to the Mr. Potato Head task, with a variety of object scales. The Bennet test eliminated the step of needing to identify the correct parts to match for each test condition that the Mr. Potato Head task had demanded – this likely eliminated a confounding variable where participants may have chosen the wrong parts. The Bennet task was more repeatable and predictable and was an efficient way to sample object interaction and tool use with objects ranging in size from a large crescent wrench to small nuts and bolts. The Bennet test seems well suited to use a generalized tool task for studies with similar goals to this one.

The O'Conner task also provided an update to the Mr. Potato Head task, with a sorting, selecting, and placing component. However, the O'Conner task tested the participant's ability to select and manipulate multiple simultaneous identical objects at a very small scale, which is set of conditions that was not tested for in the original protocol. This task provided some very valuable insight into the encumbrance effects happening at this small scale, and the increased role that sensing played in this small object manipulation task.

The Minnesota dexterity task provided a similar function to the Box and Blocks task from the first study, with one important update – each disk had an equal chance to be grasped by the participant. The blocks in the Box and Blocks task were piled on top of each other, which meant that participants would frequently struggle to select a single block that was wedged into a row of other blocks. While this is an accepted part of the Box and Blocks task protocol, it likely provided a subtle confounding variable – which is eliminated by the even spacing of the Minnesota disks. The Minnesota task was fast and reliable – fast enough that it was run twice in each study round, which allowed for repeated measures to be collected. This task seems to be an excellent assessment of repeated medium-sized object interactions at speed.

Collectively, these three new tasks provided a wide range of benefits and dealt with some of the challenges inherent to the original task set. While some of the original tasks, such as the Box and Blocks task, may still be valid measures and appropriate for inclusion in some future studies, the revised tasks of the Cost of Haptics protocol are able to efficiently measure a wide range of generalized hand interactions, and provide a range of valid metrics.

6.5.3 Tactile Discrimination test

The Tactile Discrimination test was intended to replace the Geometric solids test from the original protocol – which provided some interesting results but suffered from a high degree of variability which made the results somewhat unreliable. Using time as the metric for a shape sensing test seemed to be the cause of the variability, and so a new test was needed that used accuracy as the metric. The STI2 test looked promising but was ultimately considered to be too easy for the range of sensing capabilities being tested in this study. This meant that the Tactile Discrimination test was developed to address these issues and became the only in-house developed task included in this new protocol.

The Tactile Discrimination test showed significant difference between the gloves in Study 2 with 4 of the six shape conditions. This is an improvement over the STI2 test, which in preliminary testing was unable to show differences between the gloves at its lowest scale of 5mm. In this way, the Tactile Discrimination test was successful, but there is also room for continued refinement.

Some of the variability in the results for this test comes from the relative freedom participants are given in their approach to identifying the shapes. They are allowed to use whatever methods seem appropriate to them and aren't restricted to controlled movements or body parts. This approach was selected as the goal was to test the gloves, not a particular aspect of the human sensing apparatus, and it seemed appropriate to ask users to find their own solution to the problems the gloves created for them. If a solution they discovered was possible while wearing a specific glove, that provided some information about the limits of the glove's encumbrance. This does lead to some potential confounds, however, so future studies may wish to control the participant's touch interactions more closely.

This test did highlight the value that participants placed on using their nails to sense features that were smaller than they could detect with their fingertips alone. This suggests that fingernails may play a role in small scale sensing for MR gloves, and designs should be considered that do not block the wearer's nails.

One unresolved question with the performance of the Tactile Discrimination test is whether participants are capable of identifying the shapes directly, or whether they are relying on adjacent comparisons to make their decision. It is also not clear if there is a threshold where the comparison becomes required at a smaller scale. The STI2 test had only three options, so it was relatively simple to figure out what the remaining shape in the set was if you had a positive identification of the first two. The Tactile Discrimination test has six shapes, but as they are arranged in sets of three pairs, this process of elimination may also be a factor in participant's identification strategy.

The remote apparatus itself may also provide the user with some extra information about the state of the test. The stepper motor that drives the turntable makes a noise, and the duration of that noise may help the participant guess which position the table has turned to. It is not clear that any participant was using this noise to aid their identification, but it is hard to determine if they were doing so over the remote video stream. A possible solution for this is to introduce a random move prior to each correct move, which would obfuscate the information in the turntable noise, and make it harder to guess. The task was originally designed to be used in a lab setting, so the remote apparatus performed remarkably well, but this is an opportunity for improvement in future remote studies.

6.5.4 Remote study and participant sample

As discussed, this protocol was not originally intended to be run as a remote study, but the COVID-19 requirements made this essential. Overall, this remote study was successful – returning a huge set of data with significant results across almost all test and metrics. However, there are aspects of the remote study experience that likely influenced the results.

The need to run the study in the participant's homes meant that the there was little control over the environment of the study setting. This made it impossible to eliminate the confounds that may come from the environment the study was set in. While participants were instructed to create a space with limited distractions, the study was run over the course of three days, and there were numerous opportunities for interruptions. Participants also had varying access to space for this study in their homes, and varying sizes of tables. This meant that some participants could only have one task at a time set up on their table and required them to do a full setup and breakdown with every task switch.

The remote nature of the study meant the participant and facilitator were each reliant on a stable internet connection, which was not guaranteed by the recruitment. While the study was able to proceed, there were a number of instances where unstable internet issues halted the study and required tech support and time to address. Similarly, while there were few instances of technical breakdown with the study apparatus, the few occasions where something happened were much harder to fix, as the facilitator needed to guide the participant to do it themselves. On one occasion, this took 30 minutes to restore access to a critical camera feed and would have added to the fatigue and frustration felt by the participant.

Fatigue was generally a factor across the study, as the remote protocol required three consecutive days from the participant. The participants were required to set up, clean, and re-pack the apparatus themselves, and they were asked to perform a full task set up to six times a day. Due to the requirements of the sanitation protocol, they were also asked to participate in Study 2 and 3 concurrently, which would have added to their fatigue. While the protocol implemented counterbalancing, this may also have led to some practice effects with the number of repetitions the participants completed. These would not be expected effects in a lab setting, but they were an unfortunate requirement in this remote setting.

The sample was unusually small for this study, with six total participants recruited. This was again due to limitations imposed by the COVID-19 human subjects research policies – which greatly limited the available sample pool, while limiting the number of participants that could be run through the protocol to one a week. This small sample yielded a large data set, the large majority of which was capable of showing significant differences between the gloves. However, this small sample of research assistants offers a very narrow slice of the broader population, which may limit the wider application of these findings. This is a necessary practical consequence of conducting research during a pandemic, but there should be an opportunity in the future to confirm these findings in the lab with a more diverse sample.

6.5.5 Scaled response questions

Some of these remote protocol fatigue effects may have had an effect on the participant's results for the scaled response questions, as many of the questions returned insignificant results. One notable area where this was observed was the mental and physical load scales in the NASA-TLX instrument, which returned almost universally similar results. This is in contrast to Study 1, where the SEQ results identified reliably significant differences with the same set of gloves.

The intent to switch to the NASA-TLX tool came from observations during the early pilot for Study 2 that participants were using the SEQ occasionally as a frustration scale, rather than a difficulty scale. As previous studies have highlighted the increased cognitive load of some of the participant task adaptations, it seemed practical to split out the comprehensive difficulty of the SEQ into cognitive and physical task load, and to ask specifically about frustration. Fortunately, the NASA-TLX instrument already has these three scales developed, so it was easy to incorporate into the new protocol.

It is entirely possible that there were no confounding effects in play with the insignificant results for the mental and physical load questions. However, it does seem unlikely that the task experiences varied across the board only by frustration, and never by any significant mental or physical load. One consequence of the extra time pressures of the remote protocol is that it was no longer possible to probe more deeply with individual answers to scaled response questions. It is possible that this, in some combination with the small N, and fatigue effects, led to results for the NASA-TLX that collapsed the apparent differences in task load. This may need to be confirmed in a lab setting in the future.

One final observation about scaled response questions is that the final glove preference survey may benefit from the addition of ranked choice voting. The forced choice to select a single glove in the current version is effective at providing a clear outcome that best represents their overall choice, but a ranked choice option might provide some nuance that is currently missing. It may be valuable to compare the top choice to the second or third choice to see if they have similar characteristics. It may also be helpful to know which glove reliably is chosen last.

6.5.6 *Glove observations*

The participants in Study 2 commented on the gloves being itchy at a higher rate than previous samples – this may be responsible for some of the differences in comfort ratings. The introduction of the O'Conner dexterity test also allowed for a new effect to be observed – as seen in Figure 77, the Hockey glove's knit webbing getting pins stuck in it during the course of the task. This aligns with an observation that participants may have been using the mesh of the USG/Powermesh glove to snag on corners in the Tactile Discrimination test, which may have provided additional information. These open mesh or webbing structures in the fabric appear to have an enhanced ability to snag and catch on objects and environments, which may show effects across various tasks.

Figure 77 – Two photos showing the Hockey gloves getting pins from the O'Conner dexterity test stuck in the webbing sewn on the sides of the glove fingers.

6.5.7 Conclusion

Study 2 successfully introduced the updated Cost of Haptics protocol, and demonstrated the implementation of the remote protocol, which allowed this study to proceed during a glove pandemic. This study replicated much of the results of the previous study – showing the ongoing effects of slip/grip on the glove surfaces, the trade-offs between glove features that support sensing and manipulating, and the heavy encumbrance effects of the Hockey glove. This will be the last study to feature this set of Commercial gloves, though the USG/Powermesh glove will continue on in the next study, which focuses on smaller feature differentiation, and tests single independent variables, rather than the aggregated design variables present in the Commercial glove set.

CHAPTER 7. STUDY 3: CONSTRUCTED GLOVE STUDY

7.1 Introduction

After proving out the efficacy of the new tasks in the protocol, the next study switched to focus on the testing of discrete independent variables. This was accomplished through the creation of custom gloves, constructed to be identical, save for the variation of a single variable. By testing these gloves inside a task set, it was proposed that it would be possible to see the encumbrance effects of these specific design features.

Figure 78 – The six glove conditions for Study 3: Constructed Glove Study

This study was structured to test a set of separate independent variables within a single unified protocol. This was possible as all test conditions share the same control condition – the participant not wearing a glove – and the same null hypothesis. This allowed for a streamlined protocol that yielded data to test multiple parallel hypotheses. The gloves used for this study were constructed to order, the process of which is detailed in Chapter 6.

7.1.1 Hypotheses

The common hypothesis across all conditions were as follows:

> H_0: There will be no significant difference in performance between the glove conditions.
>
> H_1: There will be a significant difference in performance between the glove conditions.

If significant differences in the measured effects between the gloves are observed, a set of secondary hypotheses were then tested:

> H_2: The “no glove” condition will have better performance across all tasks than any glove condition.
>
> H_3: The “Basic powermesh glove + embedded VT at fingertips” condition will show significant impairment in performance compared to the “no glove” condition.
>
> H_4: Participants will express an overall preference for the “Basic powermesh glove + embedded VT at fingertips” condition.

Testing these hypotheses will yield new insight as to the effect size of the encumbrance effects for each of these specific design features and will test the method for evaluating these types of independent variables.

7.2 Methods

7.2.1 Participants

Participants for this study were recruited from a pool of Research Assistants, and were approved for participation in a remote study. Due to limitations from COVID-19 policies, the sample was capped at 6 participants (4 self-identified as female, 2 as male), with an age range between 24 and 33 years of age. As discussed in Chapter 5.4.1, the participants for this study were also selected to participate in Study 3, as both studies shared the same apparatus and sanitation protocol. While the same participants completed both studies, they were not considered as single group across both studies, and the studies were not run concurrently. Participants were required to meet the inclusion criteria identified in Chapter 5, and be willing to conduct the study remotely in their own homes. The protocol was approved by Western IRB, and the Georgia Tech IRB (in reliance on the Western IRB approval), and all participants provided their written informed consent.

7.2.2 Study Design

This study was structured to test each combination of glove condition and task, while controlling for ordering effects. This was done by using a latin square to structure the order of the glove conditions and using a random generator to structure the order of the tasks within each glove condition. Six glove conditions are included in this study, which are discussed in detail in Chapter 6:

- Glove 0: no glove

- Glove variant 1: Basic powermesh glove
- Glove variant 2: Basic glove built from a stretchy knit (Nilo)
- Glove variant 3: Basic powermesh glove + 50% coverage palmar TPU
- Glove variant 4: Basic powermesh glove + 100% coverage palmar TPU
- Glove variant 5: Basic powermesh glove + embedded VT at fingertips

7.2.3 Apparatus

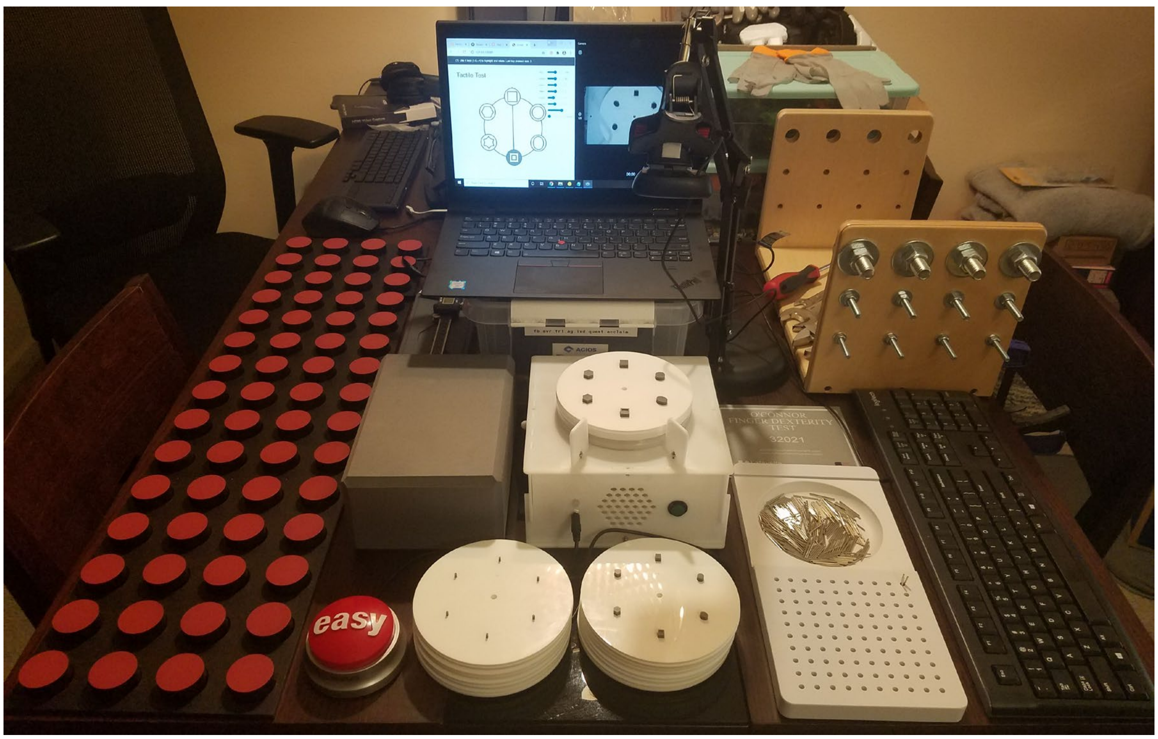


Figure 79 – The assembled apparatus for the Cost of Haptics Study

Five tasks were included in this study – shown in Figure 79 – as detailed in Chapter 4.2:

1. Typing test
2. O'Conner dexterity test

3. Minnesota dexterity test
4. Bennett hand tool dexterity test
5. Tactile Discrimination test

Each of these tasks was administered to the participant using the requisite hardware, which the participant configured under the guidance of the study facilitator. Additional apparatus included a laptop, camera equipment, and cleaning supplies, which were used to support the remote operation of the study, as detailed in Chapter 5.4.

7.2.4 Measures

Each of the five tasks had a primary measure of performance, and a set of associated subjective measures. The performance measures are as follows:

Table 8 – Performance measures for Study 3 tasks

Task	Measure	Unit	Calculation
Typing test	WPM	count	Correct, complete words / minute
	Error rate	%	Correct words / total words
Bennet hand tool test	Speed	seconds	Time to completion
	Error rate	count	Dropped parts during task
O'Conner pin test	Speed	seconds	Time to completion
	Error rate	count	Dropped pins during task
Minnesota turning test	Speed	seconds	Time to completion
Tactile discrimination test	Scale	mm	Smallest scale correctly identified

The subjective measures for this study are discussed in depth in Chapter 4.3. The set included for the study are as follows:

- NASA-TLX Scale – Mental load, Physical load, Frustration [0-20]
- Scaled response typing questions [1-5]
- Scaled response glove experience questions [1-5]
- Glove preference questions [choice of preferred glove]

7.2.5 Analysis

Analysis for this study was completed using SPSS for Windows. Performance data for these gloves and scaled response questions are considered to be non-parametric, therefore the Kruskal-Wallace test was selected to determine if there was variance between the median values for each group. The results are rendered in boxplots, which depict the median, quartiles, and outliers. SPSS renders the box as representing the interquartile range (IQ) of the central 50% of the values (ranging from the 1st to 3rd quartile), with the median indicated as a line across the box. The whisker lines show values that extend from the limits of the box to the highest and lowest values that do not exceed 1.5 times the range of the IQ. Circles and crosses indicates outliers, the latter of which represent extreme outliers with values that exceed 3 times the range of the IQ [42].

7.3 Results

7.3.1 Typing test

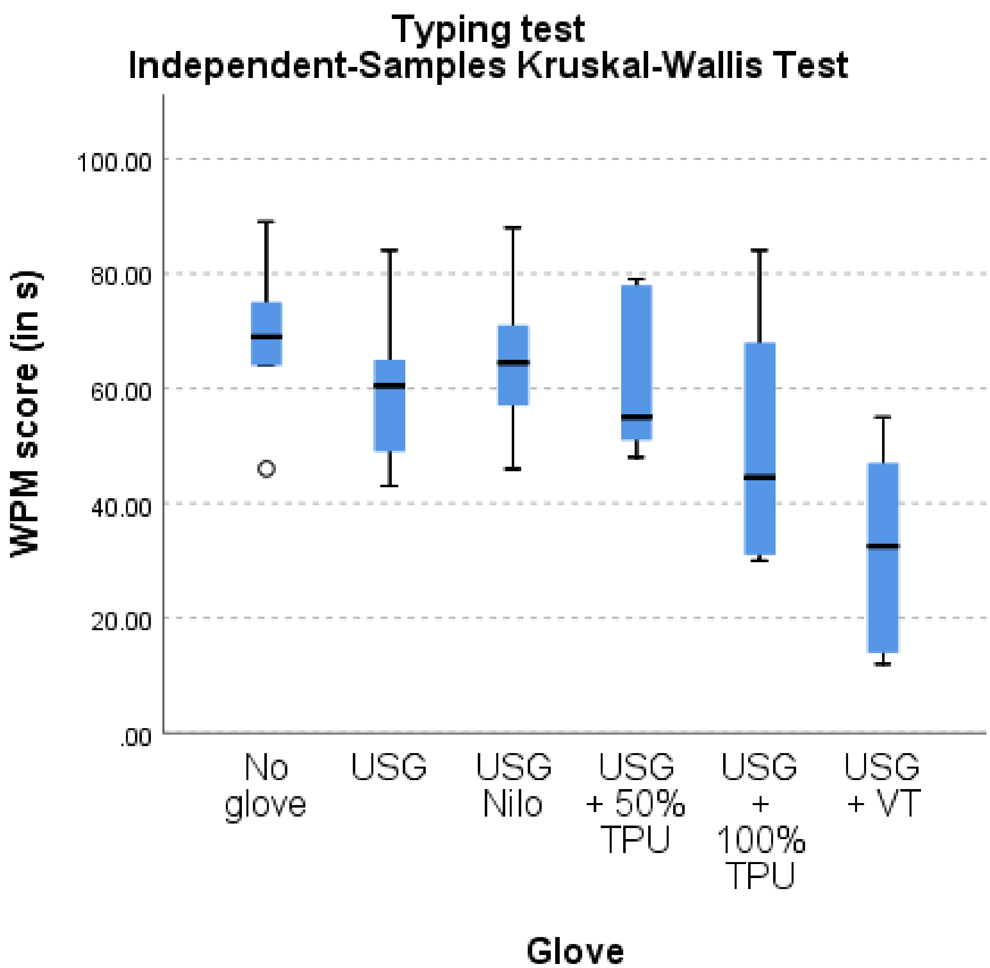


Figure 80 – Typing test Kruskal-Wallis test results – there is a significant difference in the WPM score between gloves, $\chi2(5) = 12.3$, $p = 0.031$

The Typing test results show a significant difference in WPM typing speed between gloves, shown in Figure 80. The median "no glove" WPM score was ~70 WPM, while the USG + VT glove condition showed a median score of ~30 WPM.

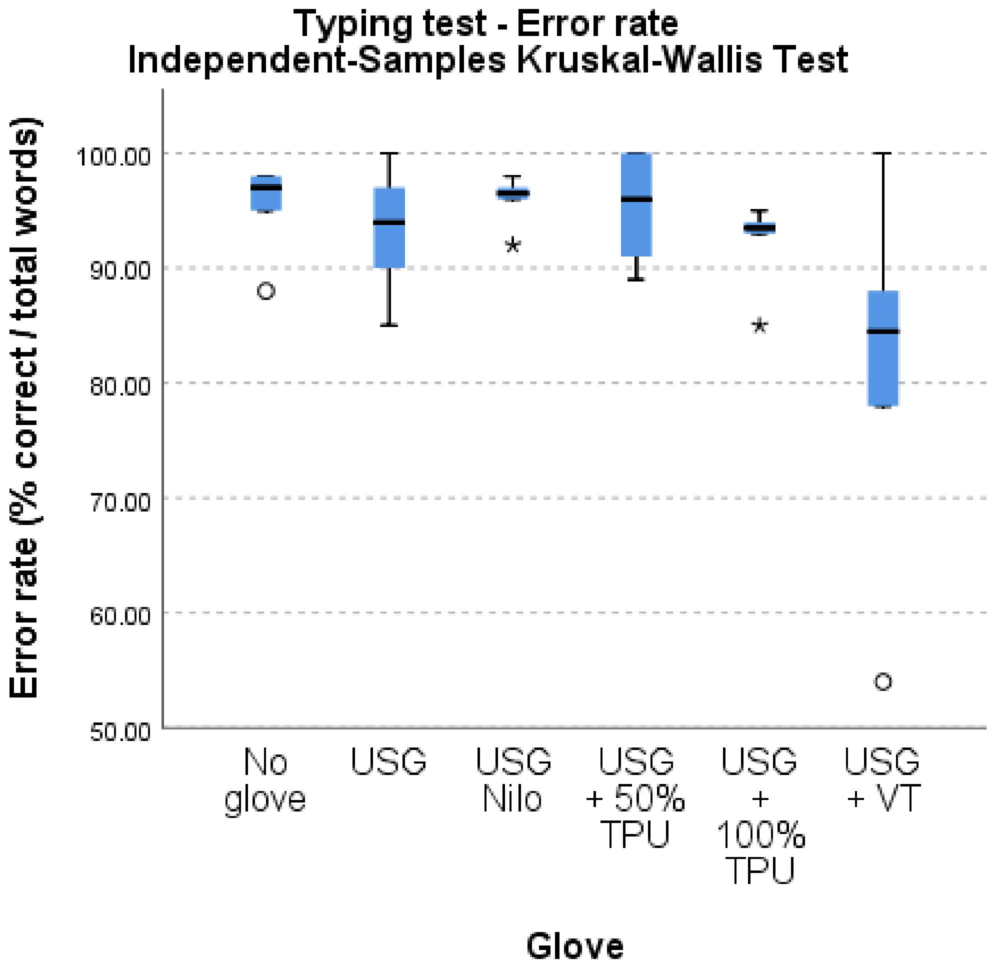


Figure 81 – Typing test error rate Kruskal-Wallis test results – there is no significant difference in the error rate between gloves, $\chi2(5) = 9.74$, $p = 0.083$

While the participant's performance showed significant differences, there was no significant difference in their Typing test error rate, as shown in Figure 81. This is primarily due to the gloves, with the exception of the USG + VT glove, all maintaining a median error rate close to ~95%, which is quite similar to the "no glove" results of ~97%. The USG + VT glove, by contrast, showed a median error rate of ~85%, with scores stretching down to ~55%.

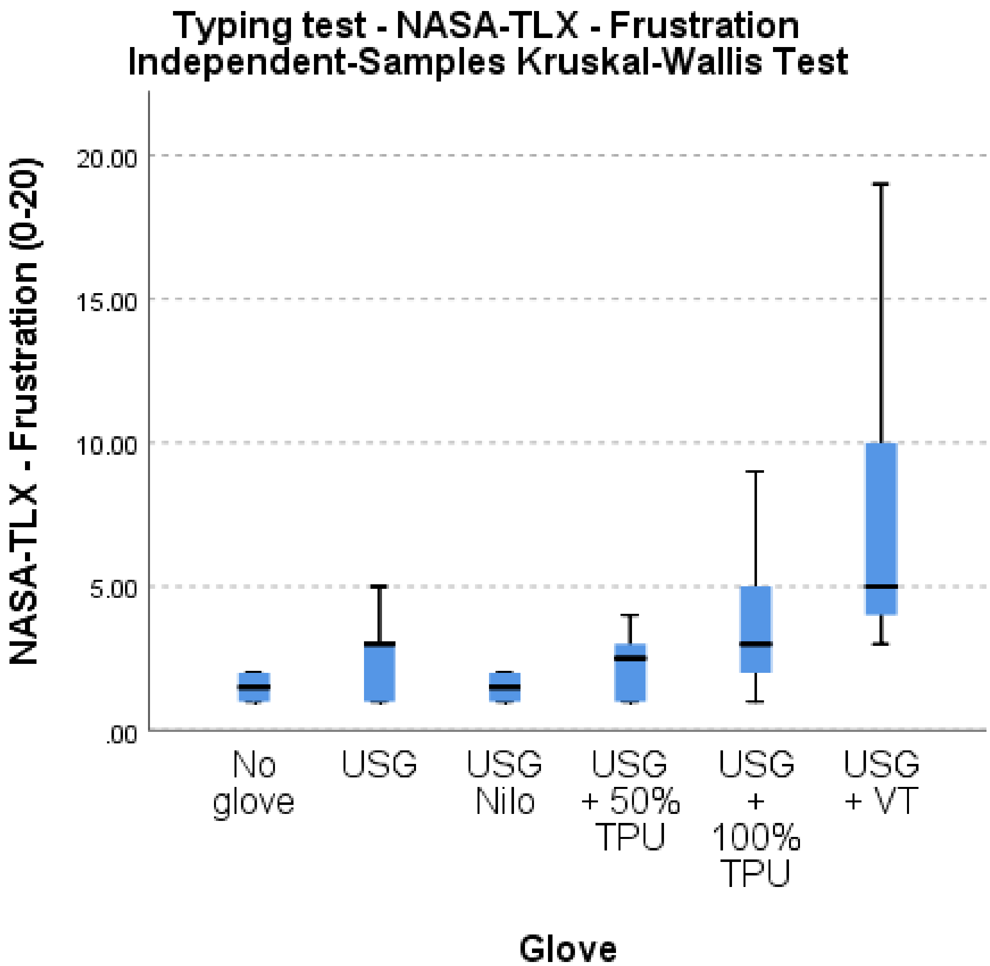


Figure 82 – Typing test NASA-TLX Frustration scale Kruskal-Wallis test results – there is a significant difference in the Frustration score between gloves, $\chi 2(5) = 15.95$, $p = 0.007$

Participant's reported a significant difference in perceived frustration in the Typing test between gloves, as shown in Figure 82. Most of the gloves scored similarly, but the significant result is due largely to the heightened frustration with the USG + VT glove. Participants reported that the VT motors made their fingertips useless, and caused them to try to type with the sides of their fingers. The motors were also reported to slip around, and prevent the participants from identifying whether their fingers were placed on top of one or multiple keys. The VTs were reported to add noticeable weight, which required the participants to adjust their typing strategy, and cause them to think more about the typing experience than they otherwise would.

7.3.2 Bennett hand tool test

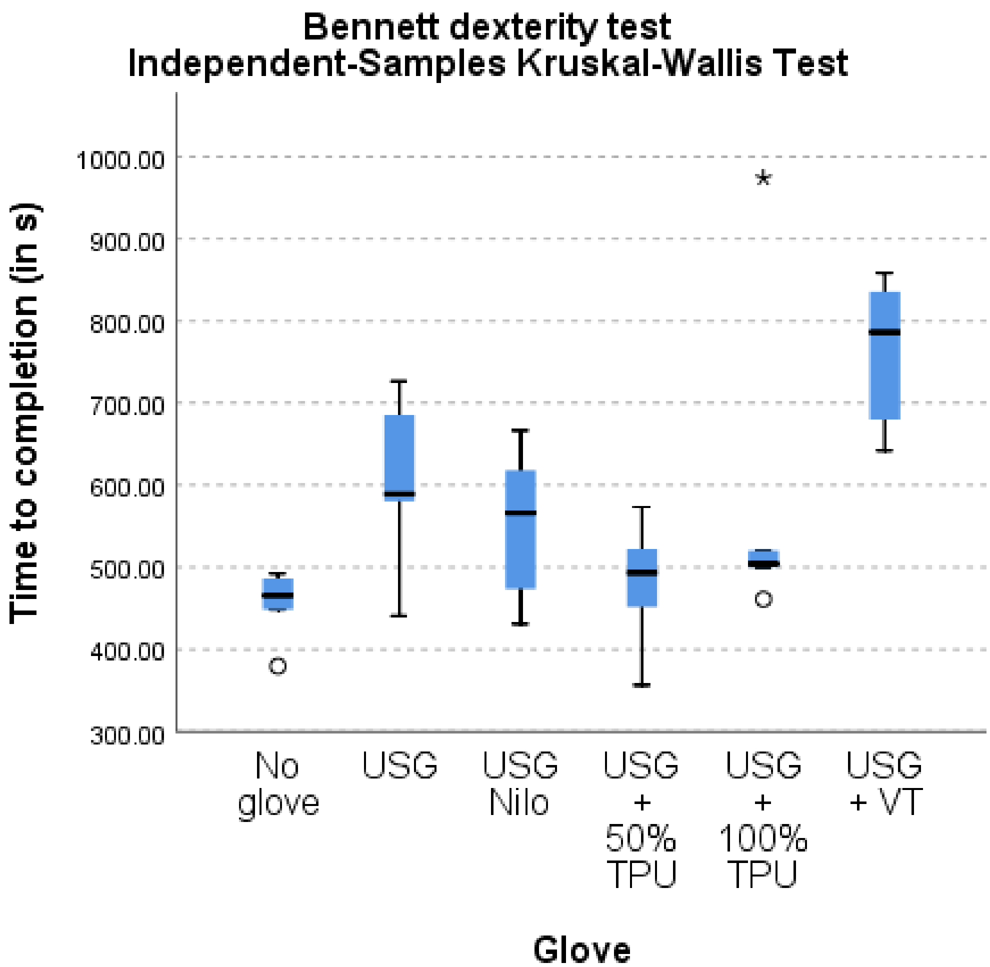


Figure 83 – Bennett dexterity test Kruskal-Wallis test results – there is a significant difference in the time to completion score between gloves, $\chi2(5) = 18.37$, $p = 0.003$

The Bennet dexterity test result showed a significant difference between gloves – seen in Figure 83. This tool manipulation test favored the TPU laminated gloves, which had median time scores close to 500 seconds, while the USG + VT glove returned a median score of 800 seconds. Participants attributed the advantage of the TPU gloves to their enhanced grip, which was useful to maintain control of the small parts. Many participants preferred the TPU gloves to the base USG, as they felt the mesh was slippery when interacting with tools.

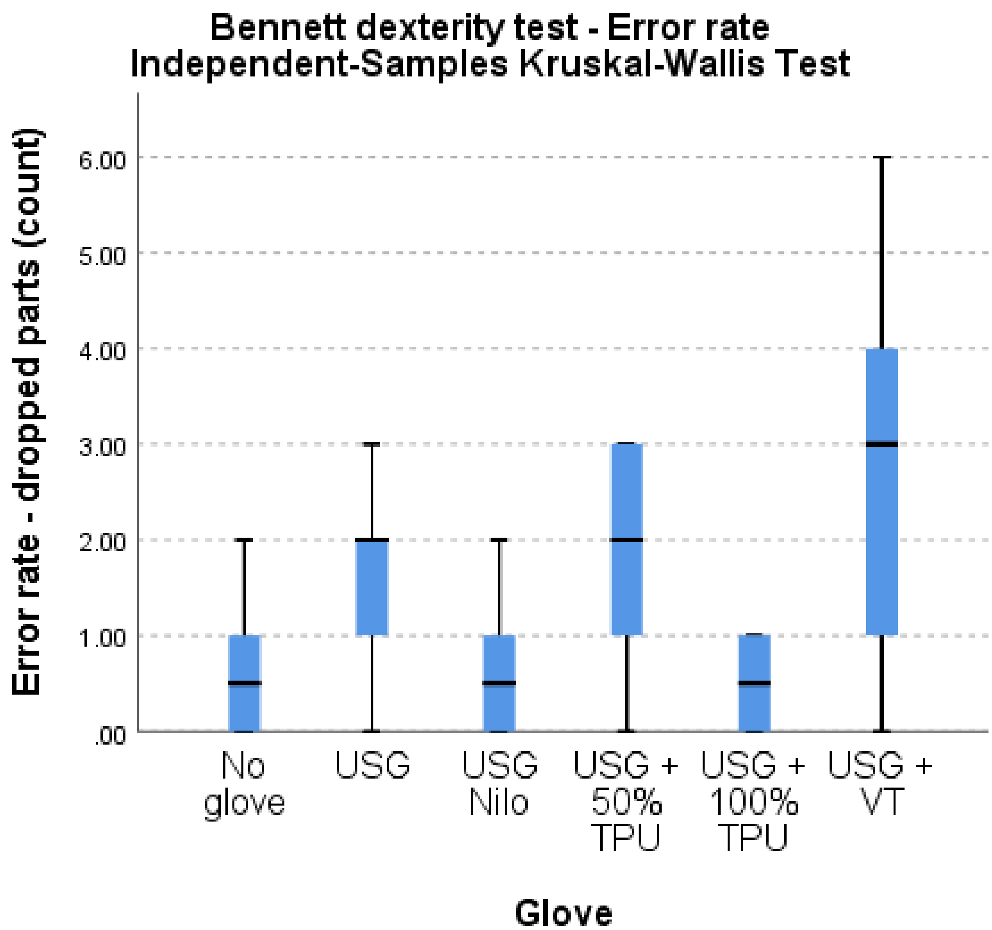


Figure 84 – Bennett dexterity test error rate Kruskal-Wallis test results – there is a significant difference in the error rate between gloves, $\chi2(5) = 11.37$, $p = 0.045$

The Bennett dexterity test showed a significant difference between gloves in the error rate – observed by counting the number of dropped tools and parts in the completion of the task. Overall, the error rate was low, with all gloves showing a median rate of three dropped parts per task or less – with the exception of the USG + VT gloves. The VT gloves showed a median error rate of three dropped parts, but had a highly variable result with one participant dropping six parts. Participant's attributed this difference to the extra weight on the fingertips, and the shifting motors causing them to lose their grip.

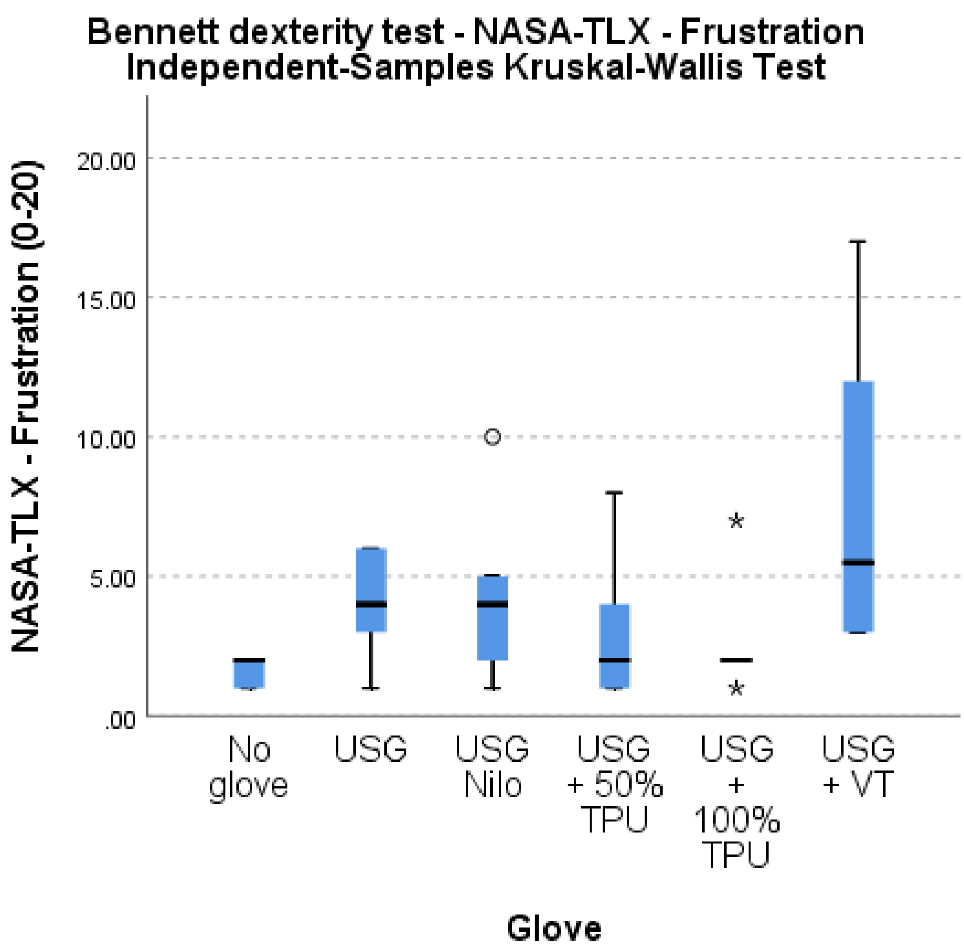


Figure 85 – Bennett dexterity test NASA-TLX Frustration scale Kruskal-Wallis test results – there is a significant difference in the Frustration scale between gloves, $\chi2(5) = 11.57$, $p = 0.041$

This result, seen in Figure 85, again matches the perceived frustration score, which showed a significant difference between gloves, again based on the poor performance of the USG + VT gloves. Participants reported that the VTs made it very difficult to adjust the wrench, much harder to spin the small nuts into place, and made it very difficult to pick up small parts off a flat surface. It is clear that the shifting motors interfered with numerous aspects of this task, and led to the significant increase in frustration.

7.3.3 O'Conner Test

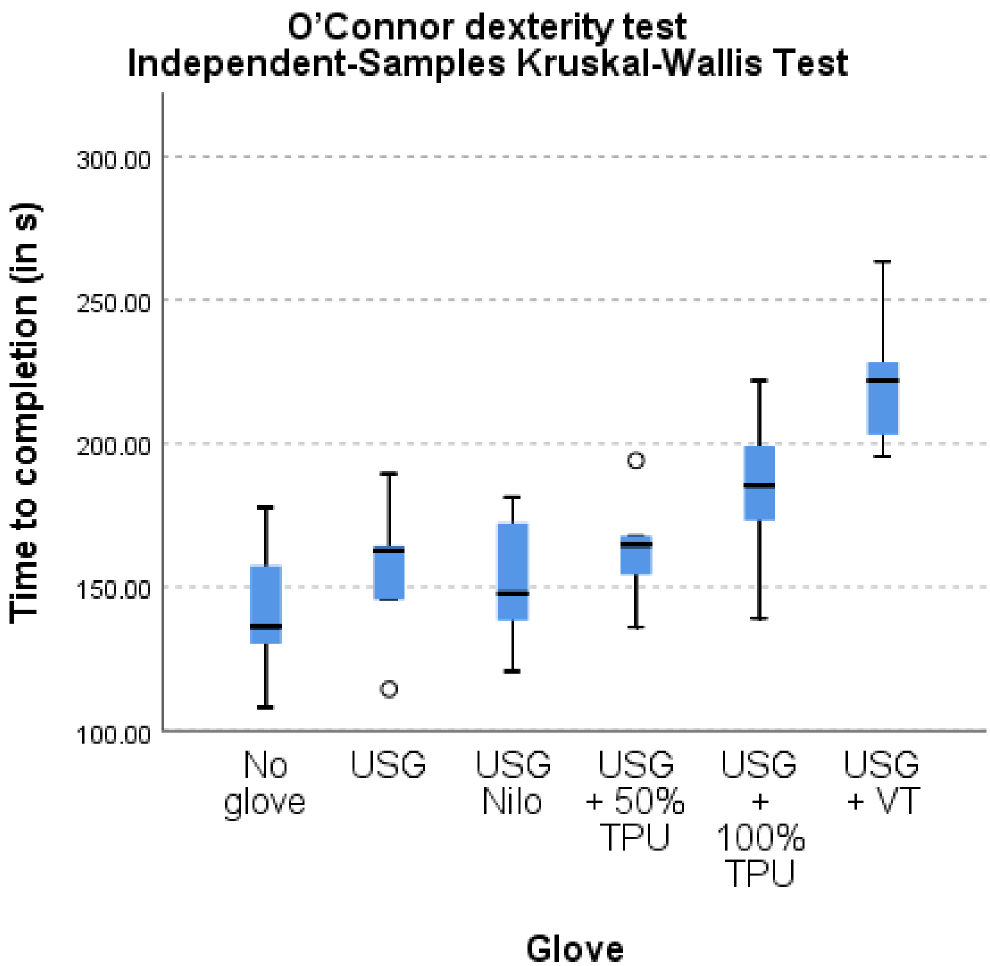


Figure 86 – O'Conner dexterity test Kruskal-Wallis test results – there is a significant difference in the time to completion score between gloves, $\chi2(5) = 19.87$, $p = 0.001$

The O'Conner dexterity test showed significant difference in results between the gloves in task performance, shown in Figure 86. The best performing glove was the USG Nilo glove, which returned a median time of ~150 seconds, while the USG + VT showed a median time of ~225 seconds to complete the task. Participants reported they were able to discern the feel of three individual pins in their fingers with the USG Nilo glove, while they reported needing to shift the pins to the sides of their fingers to feel them with the USG + VT gloves.

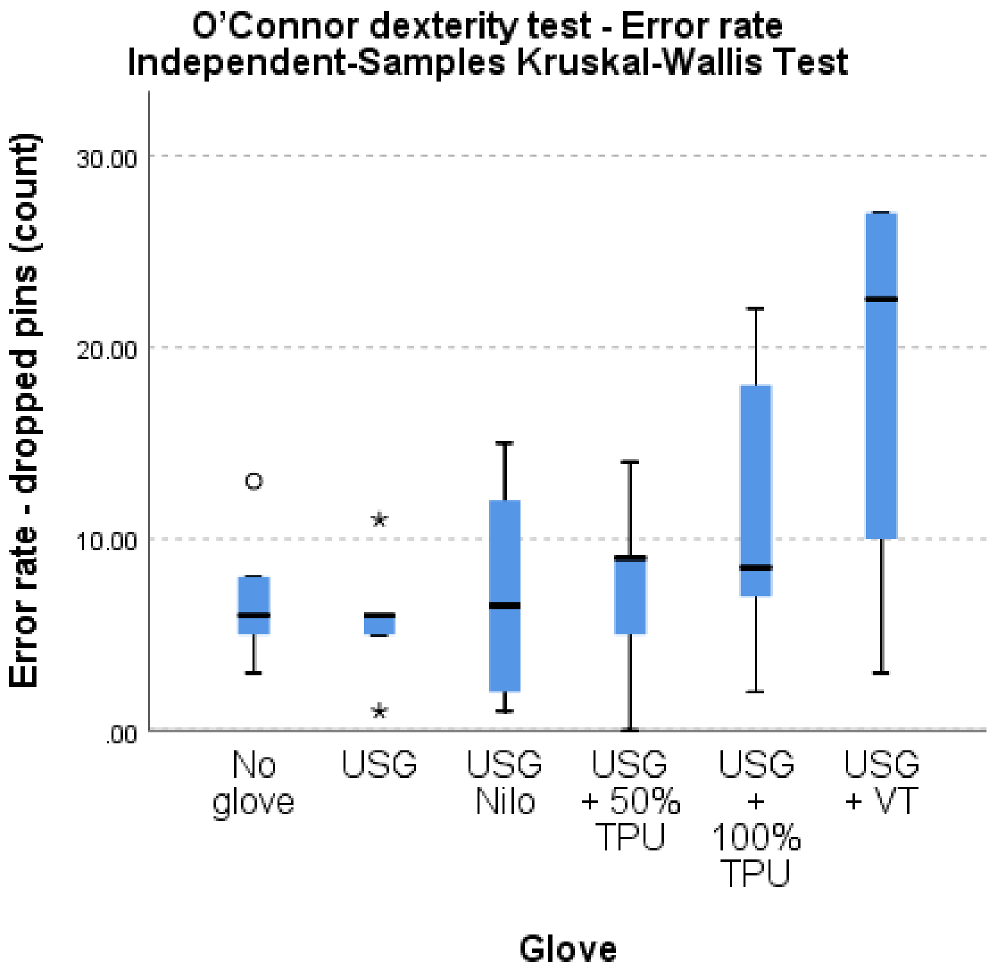


Figure 87 – O'Conner dexterity test error rate Kruskal-Wallis test results – there is no significant difference in the error rate between gloves, $\chi2(5) = 7.70$, $p = 0.174$

The error rate for this task – calculated as the number of pins dropped in the course of completing the task – showed in Figure 87 that there was no significant difference between gloves. Participants felt there were issues with all of the gloves – slipperiness with the USG and USG Nilo gloves, working around the motors in the USG + VT gloves, but they were not as prone to errors as the gloves in Study 2.

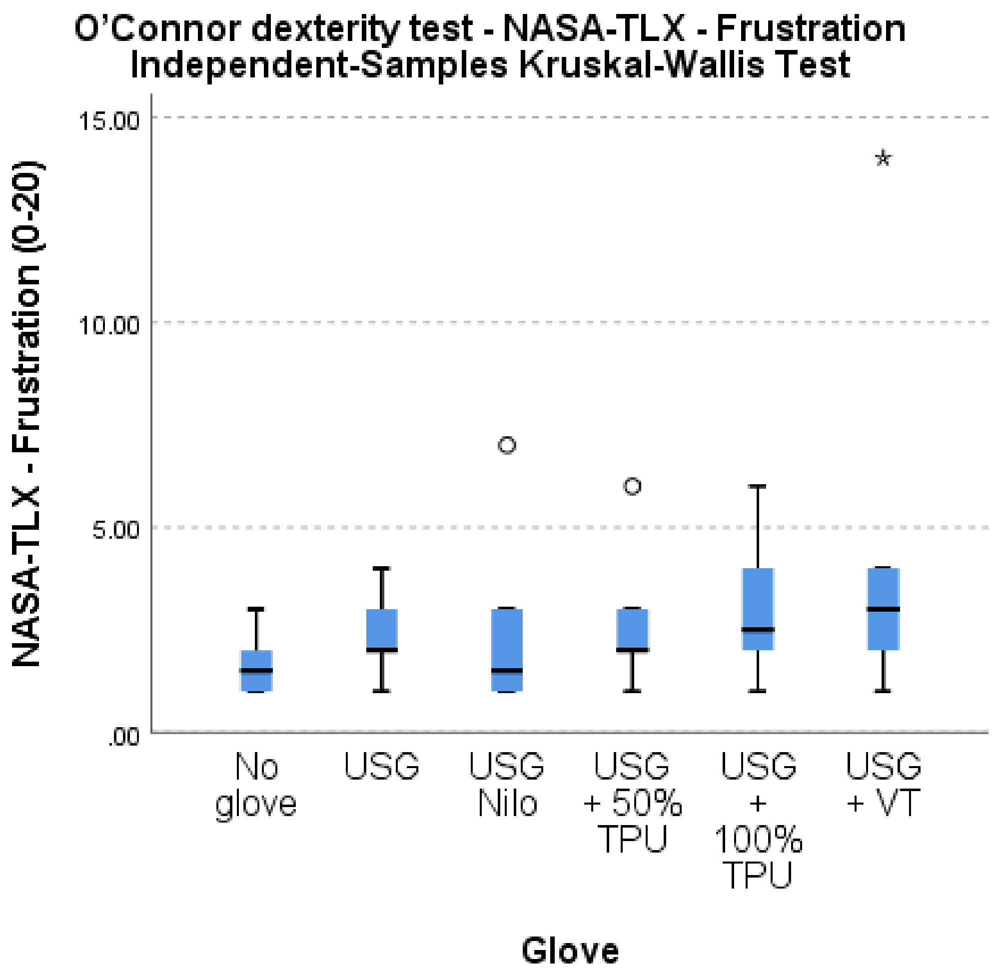


Figure 88 – O'Conner dexterity test NASA-TLX Frustration scale Kruskal-Wallis test results – there is no significant difference in the Frustration scale between gloves, $\chi2(5) = 4.48$, p = 0.483

While there were significant differences in performance between the gloves, participants did not report a significant difference in perceived frustration, shown in Figure 88. As reported, participants felt there were minor issues with each of the gloves, and had found workable strategies to deal with them. For the USG + VT, this involved shifting the motor out of the way and using the sides of their fingers.

7.3.4 Minnesota test

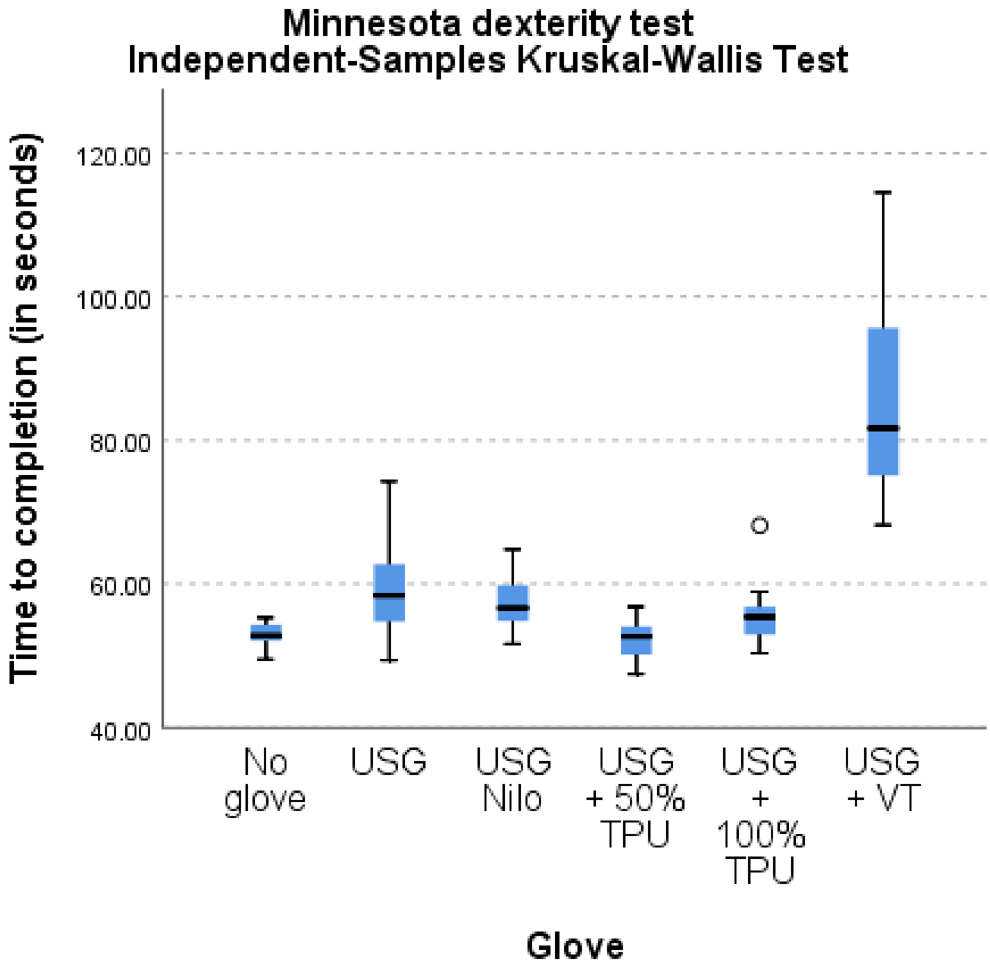


Figure 89 – Minnesota dexterity test Kruskal-Wallis test results – there is a significant difference in the time to completion score between gloves, $\chi2(5) = 42.8$, $p = 0.000$

The Minnesota dexterity test, with results seen in Figure 89, showed a significant difference between gloves in the time it took to complete the task. The USG + 50% TPU gloves and the "no glove" condition showed nearly identical results with median times at just over 50 seconds. This compares with the USG + VT gloves which had a median time over 80 seconds, with some participants needing over 100 seconds. Participants reports issues with the USG + VT glove in maintaining control over the disks and needing to squeeze harder to compensate.

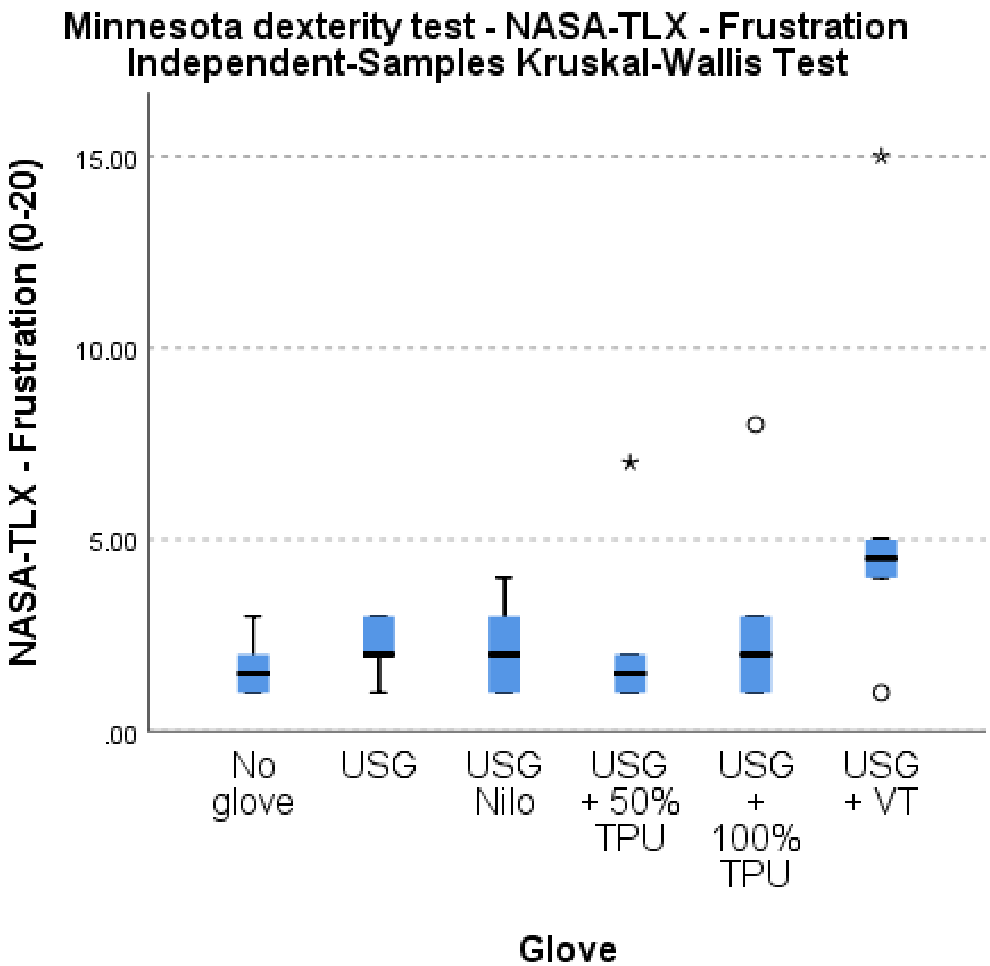


Figure 90 – Minnesota dexterity test NASA-TLX Frustration scale Kruskal-Wallis test results – there is no significant difference in the Frustration scale between gloves, $\chi 2(5) = 7.30$, $p = 0.202$

In spite of this difference, participants showed no significant difference between gloves in their perceived frustration with the task, seen in Figure 90.

7.3.5 Tactile Discrimination test

Table 9 – Hypothesis test summary for Kruskal-Wallis test across all six Tactile discrimination test conditions. Significance level is 0.05.

Task Condition	Test Statistic	Sig.
(A) Circle	7.1	.213
(B) Oval	6.39	.270
(C) Hollow Square	7.60	.180
(D) Star	6.63	.249
(E) Hexagon	1.825	.873
(F) Square	9.2	.101

The Tactile Discrimination test showed no significant differences between gloves in any of the test conditions (Table 9). This is likely due to the similarity of the glove designs, as the single change in features between glove variants appeared to have less influence on performance in this task, as compared to other tasks in the study. This is likely also attributable to the participant's adaptation strategy of moving the motor out of the way in the pocket on the USG + VT glove. This allowed participants to better expose their fingertip, which diminished the effect that the VT would have on their sensing capabilities. Participants also reported a reliance on their fingernails to detect small features, which the VT and TPU variants did not block.

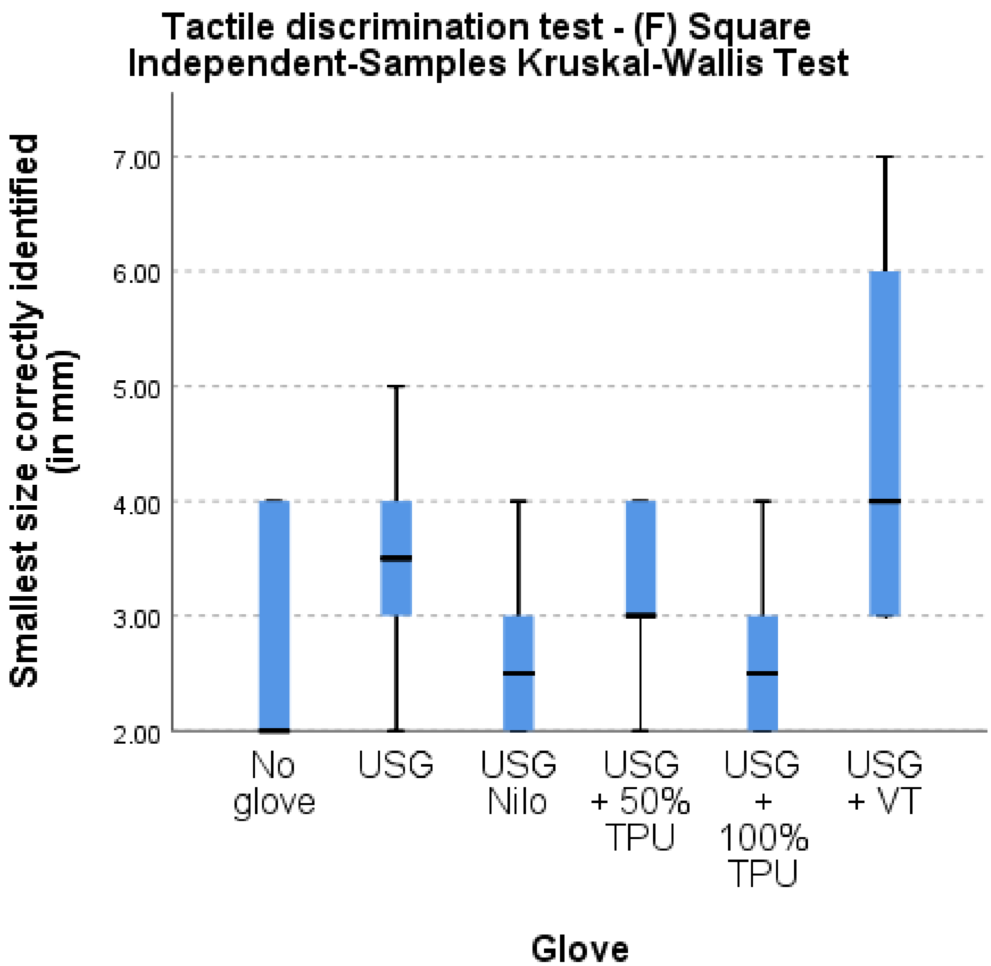


Figure 91 – Tactile discrimination test Kruskal-Wallis test results – there is no significant difference in the scale of the smallest size of the Square shape correctly identified between gloves, $\chi 2(5) = 9.2$, $p = 0.101$

While the gloves did not show a significant difference between them in the scale of the shape the participants could correctly identify, there was still notable variability in the results for the USG + VT glove, as shown in the result for the Square task condition in Figure 91. While the median score of 4mm was similar in result to the other conditions, the USG + VT gloves were the only ones where participants scored at the 6mm and 7mm scales, indicating that some of the participants experienced adverse effects.

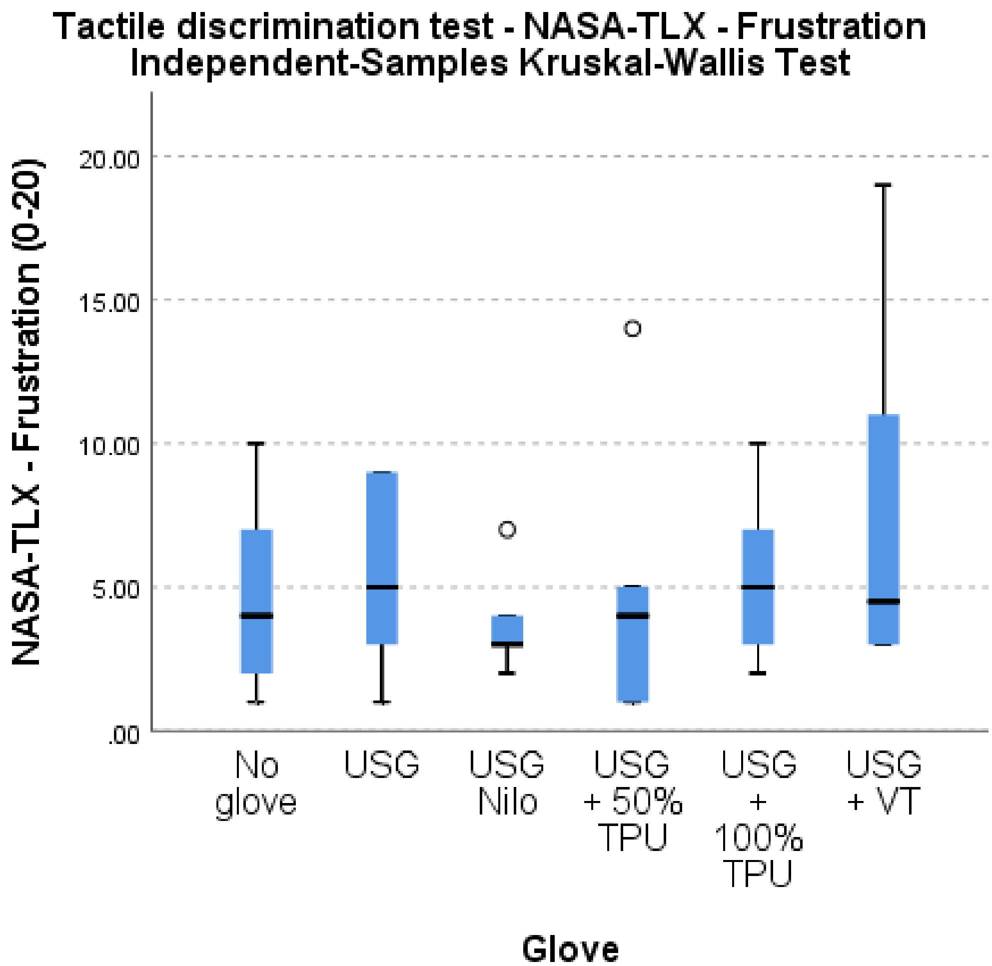


Figure 92 – Tactile discrimination test NASA-TLX Frustration scale Kruskal-Wallis test results – there is no significant difference in the Frustration scale between gloves, $\chi 2(5) = 2.43$, $p = 0.787$

The similarities continued in the perceived frustration score, which showed in Figure 92 that there was no significant difference between gloves. Participants generally found the task to be less frustrating, though the variability of the USG + VT glove is again notable, as some of the participants experienced increased frustration while wearing those gloves. The low perceived frustration score is likely again due the gloves not blocking access to the participant's fingernails or the tips of the finger in any significant way.

7.4 Glove findings

7.4.1 Scaled response – Typing questions

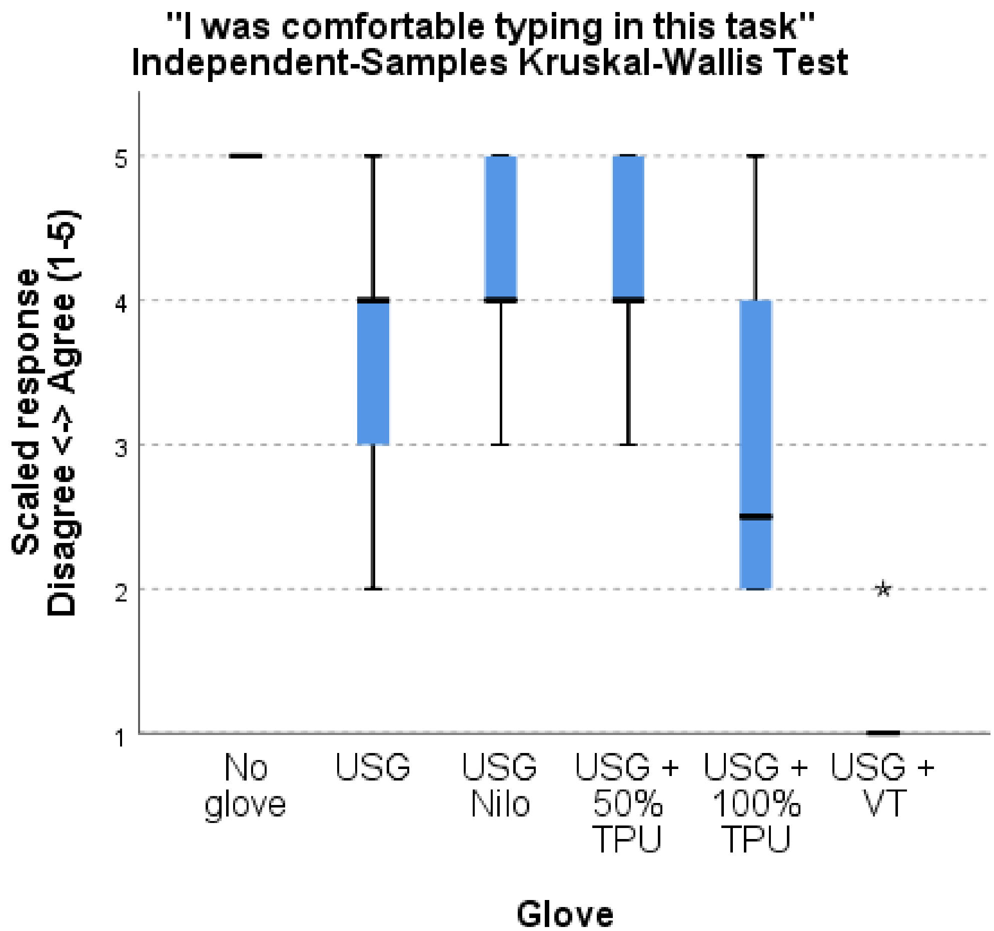


Figure 93 – Scaled response for "I was comfortable typing in this task" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi2(5) = 21.44$, $p = 0.001$

Study 3 showed significant difference in results between gloves in the effect on participant's typing experiences, shown in Figure 93. The USG Nilo and USG + 50% TPU variants both had the least detrimental impact on participant's comfort. It is notable that the USG, which is very similar to the Powermesh variant in Studies 1 & 2, did not score as well. USG + VT scored very poorly, due to their increased weight at the fingertips.

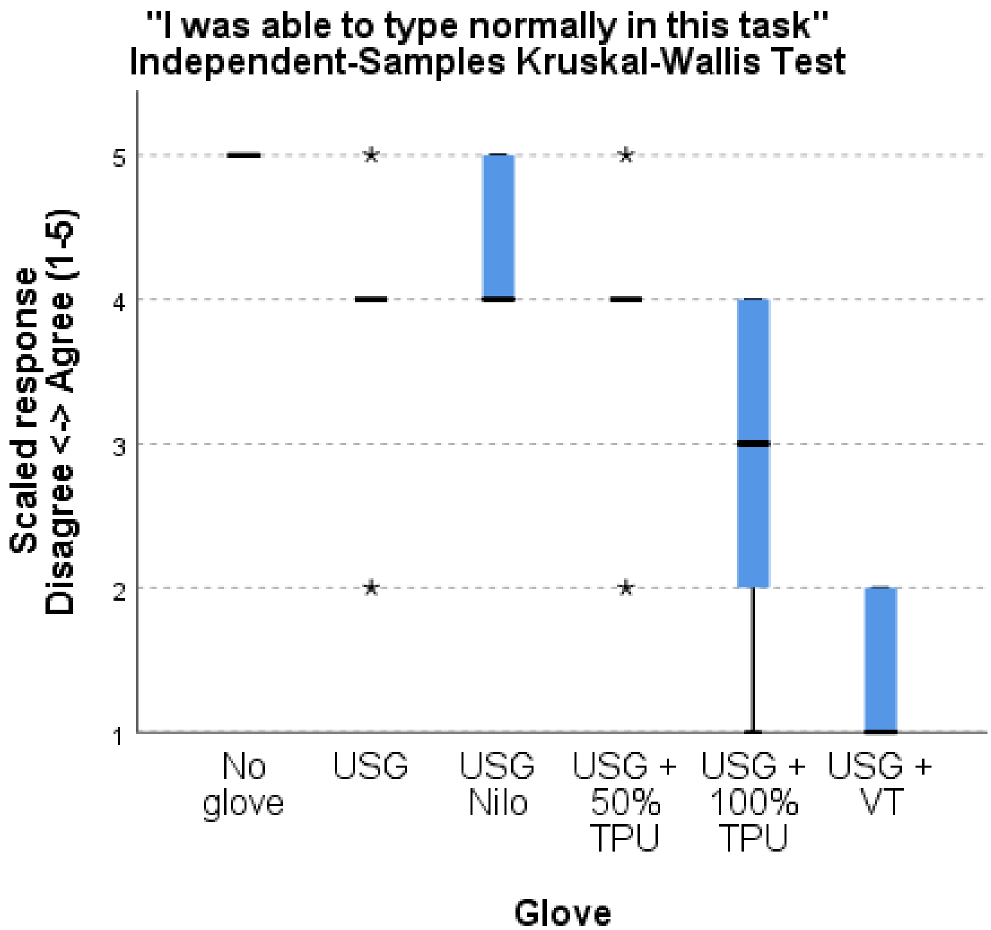


Figure 94 – Scaled response for "I was able to type normally in this task" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi 2(5) = 23.23$, p = 0.000

Results for the second question showed that the gloves were significantly different in their effect on the participant's ability to type normally, as seen in Figure 94. Participant's reported that the VT glove variant would not allow them to type normally, due to the motor obscuring their ability to sense features on the keyboard, and to discern whether they were pressing multiple keys.

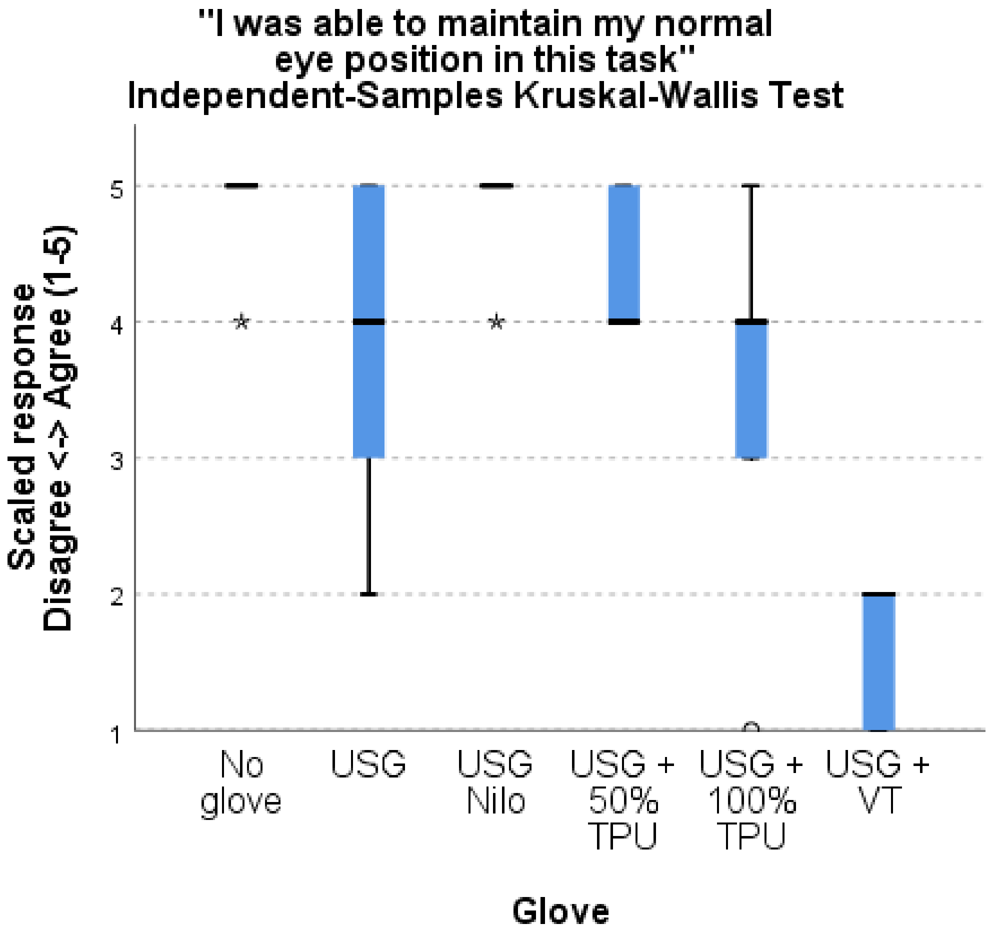


Figure 95 – Scaled response for "I was able to maintain my normal eye position in this task" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi2(5) = 19.55$, p = 0.002

Follow earlier results in the previous studies, this impairment to normal typing ability led to a corresponding significant difference in the glove's impact on eye position, shown in Figure 95. Participants reported that if they kept their gaze on the screen while wearing the USG + VT gloves, that their perceived error rate would go up sharply. The felt compelled to look at their fingers to confirm whether they were hitting the correct keys, which then slowed them down further.

7.4.2 Scaled response – Comfort and fit

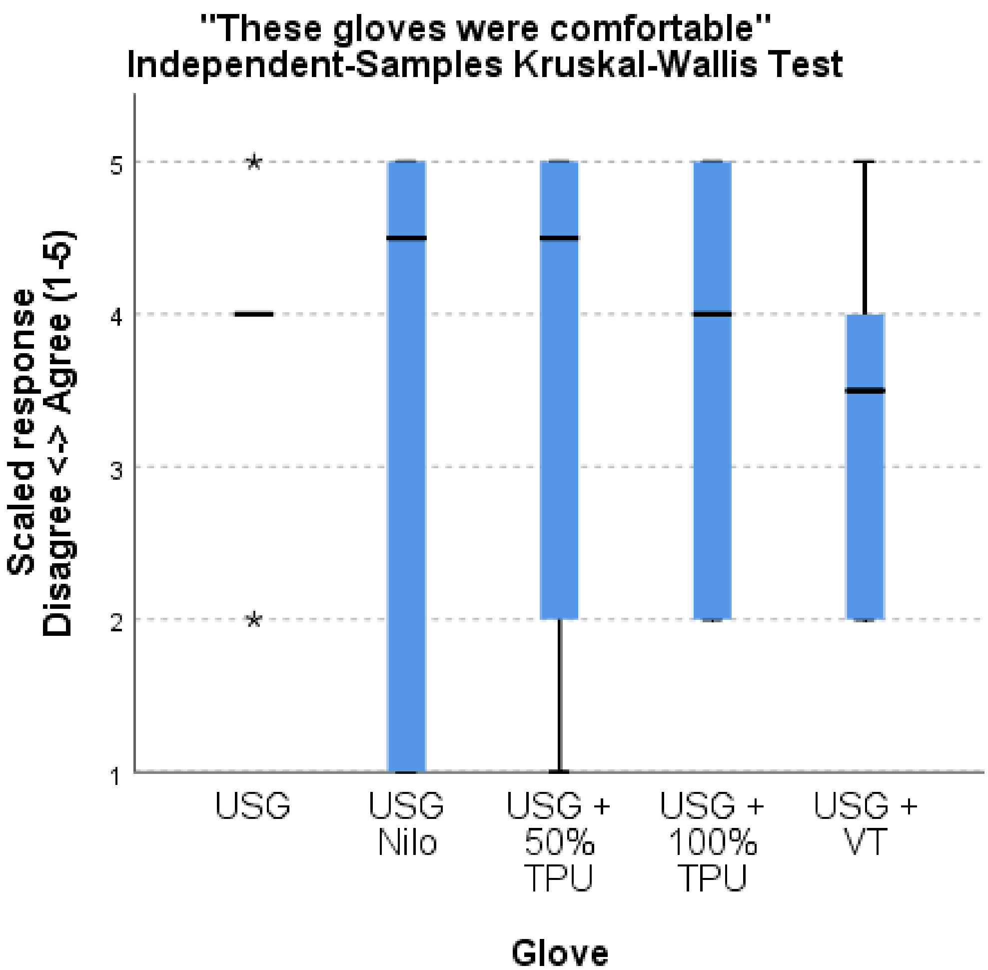


Figure 96 – Scaled response for "These gloves were comfortable" Kruskal-Wallis test results – there is no significant difference in agreement with the statement between gloves, $\chi2(4) = 0.592$, $p = 0.964$

There were no significant differences detected between gloves for questions that related to comfort, fit, and awareness of seams in the construction. This is a somewhat expected outcome for comfort, with results shown in Figure 96, as the gloves were constructed of near-identical materials, using an identical pattern. It is notable that while the median scores for this question are quite similar, there is a high degree of variability across all gloves, indicating that the participants had very different perceptions of the overall comfort of the gloves. None of the independent variables in the variants seemed to pull the result in any significant direction.

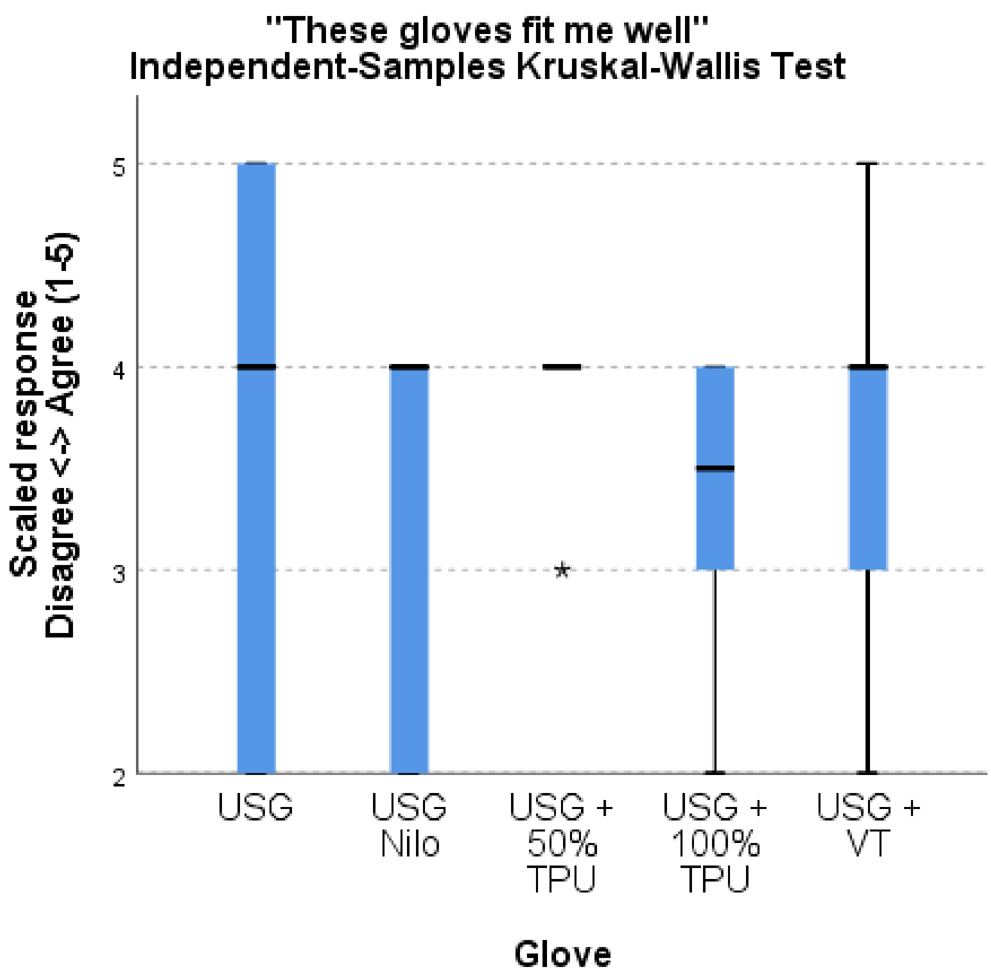


Figure 97 – Scaled response for "These gloves fit me well" Kruskal-Wallis test results – there is no significant difference in agreement with the statement between gloves, $\chi2(4) = 1.53$, $p = 0.821$

The same outcome is true for the question of fit – seen in Figure 97. Each of the gloves was built on the same pattern, and it does not appear as if any of the additional design features made a significant alteration to the perceived fit.

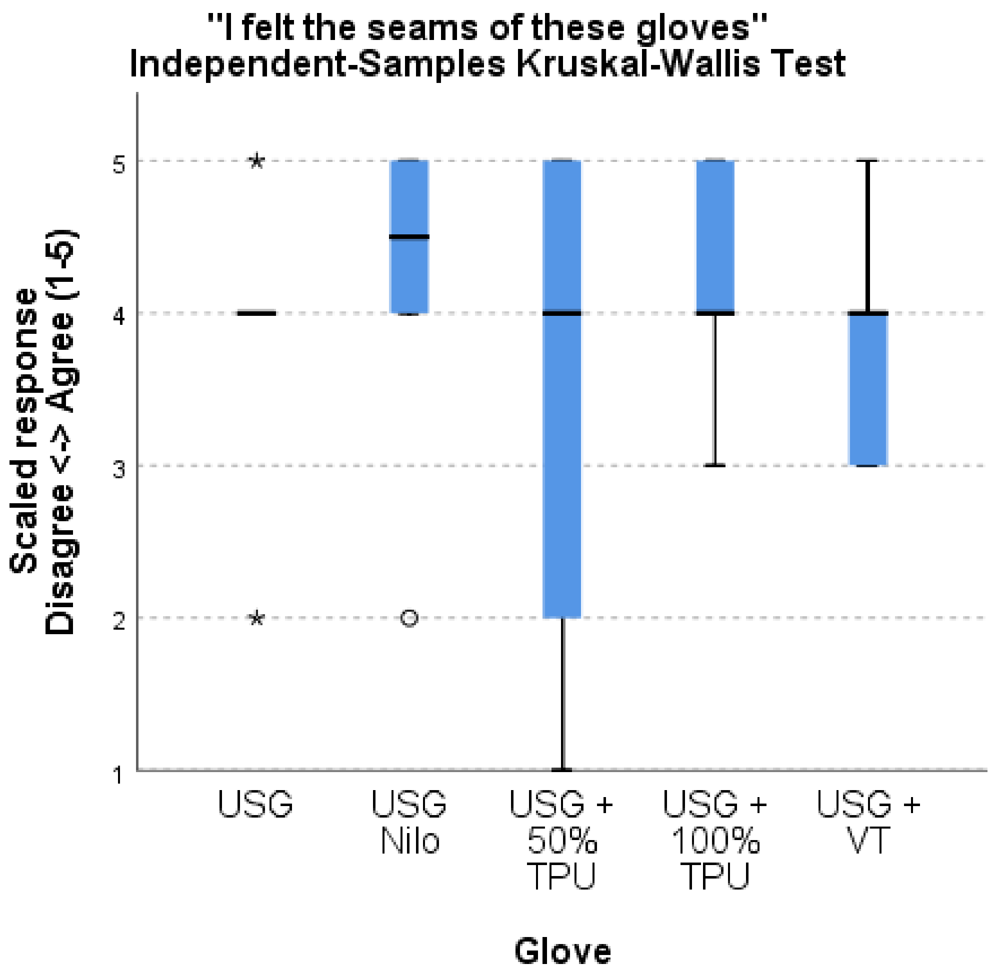


Figure 98 – Scaled response for "I felt the seams of these gloves" Kruskal-Wallis test results – there is no significant difference in agreement with the statement between gloves, $\chi2(4) = 1.47$, $p = 0.833$

The final question concerning seams returns the same insignificantly different result. The gloves were built using the same cut and sew construction method on the same graded pattern, so it is reasonable that they would have a similar median score for detected seams.

7.4.3 Scaled response – Slip and grip

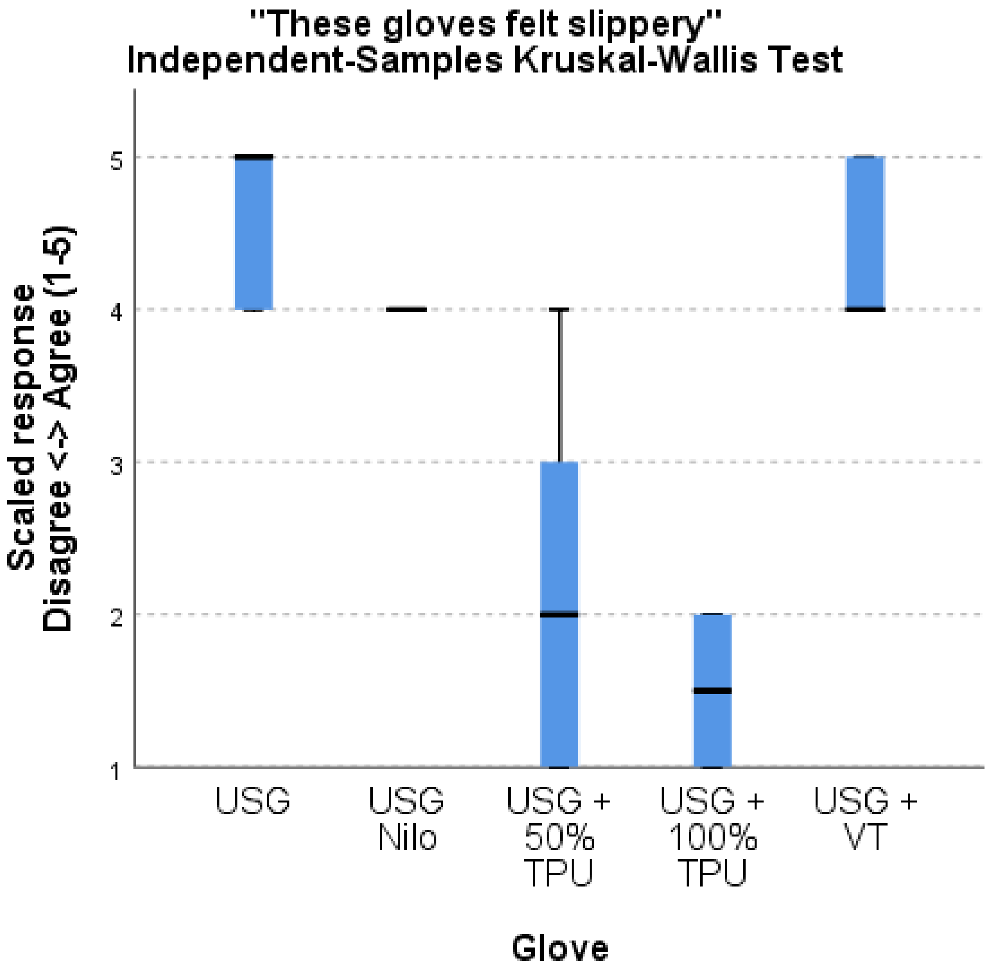


Figure 99 – Scaled response for "These gloves felt slippery" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi 2(4) = 22.77$, $p = 0.000$

The results for the slip/grip questions in Study 3 replicate the results from the previous studies. In the question of perceived slipperiness of the glove surface, there a significant differences between the gloves – with results shown in Figure 99 highlighting the slippery finish of the USG. This continues, with the USG + VT, which is also said to have a mechanical slip, as the motor moves within the containing pocket. By contrast the two TPU laminated gloves are seen to be not slippery, though there is some variability with the USG + 50% TPU variant.

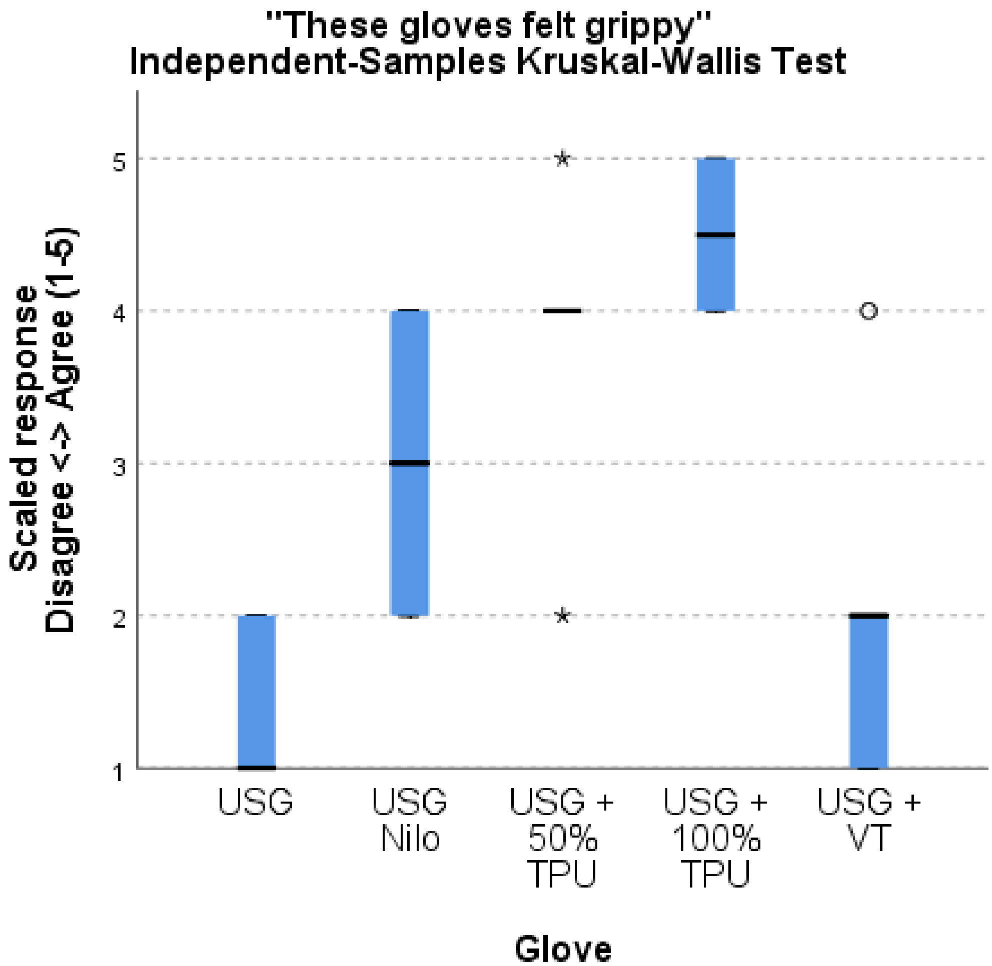


Figure 100 – Scaled response for "These gloves felt grippy" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi2(4) = 20.09$, $p = 0.000$

There is a corresponding significant result for the perceived grip capability of the gloves shown in Figure 100. The TPU glove variants are again perceived to have the highest grip, while the base USG and USG + VT is seen to have the least grip.

7.4.4 Scaled response – Hand mobility

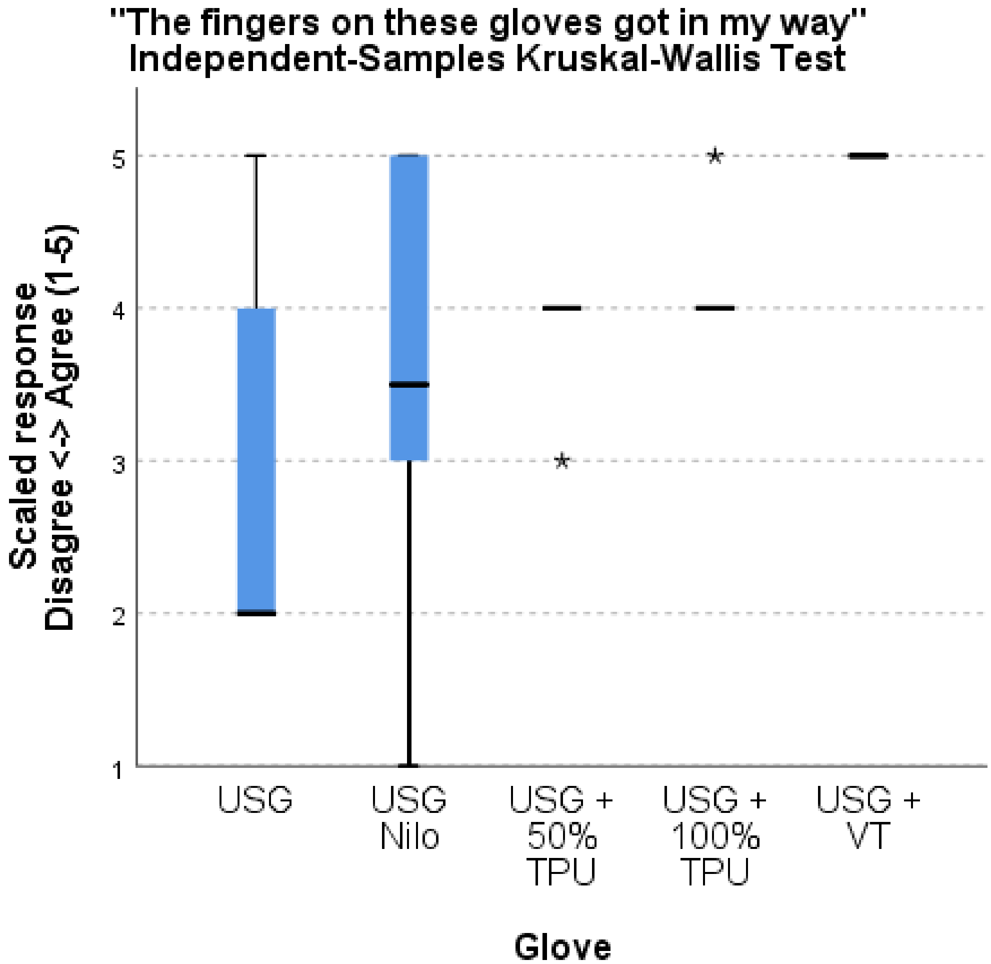


Figure 101 – Scaled response for "The fingers on these gloves got in my way" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi 2(4) = 13.16$, $p = 0.011$

The final set of questions for Study 3 relate to the glove's impact on hand mobility and physical encumbrance. As shown in Figure 101, there is a significant difference between gloves on their ability to get in the participant's way. The USG + VT scores the worst on this question, with the participants reporting that the VT inhibited both their sensing and manipulating abilities.

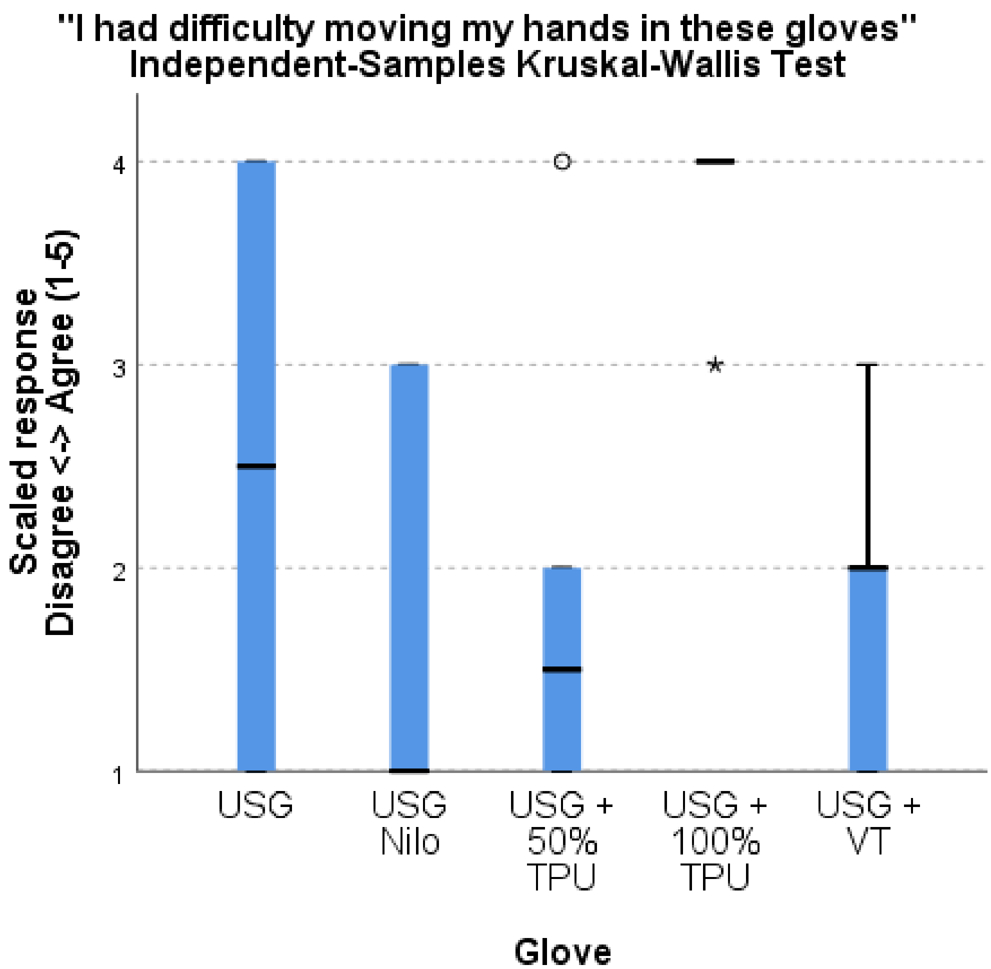


Figure 102 – Scaled response for "I had difficulty moving my hands in these gloves" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi2(4) = 11.94$, $p = 0.018$

There is a significant difference between gloves for the final question, shown in Figure 102, relating to hand movement. The notable poor performer in this question is the USG + 100% TPU variant, which has stiff TPU sheets laminated across the fingers. Participants reported that these constricted movement, especially with finger flexion.

7.5 Discussion

7.5.1 Glove preference

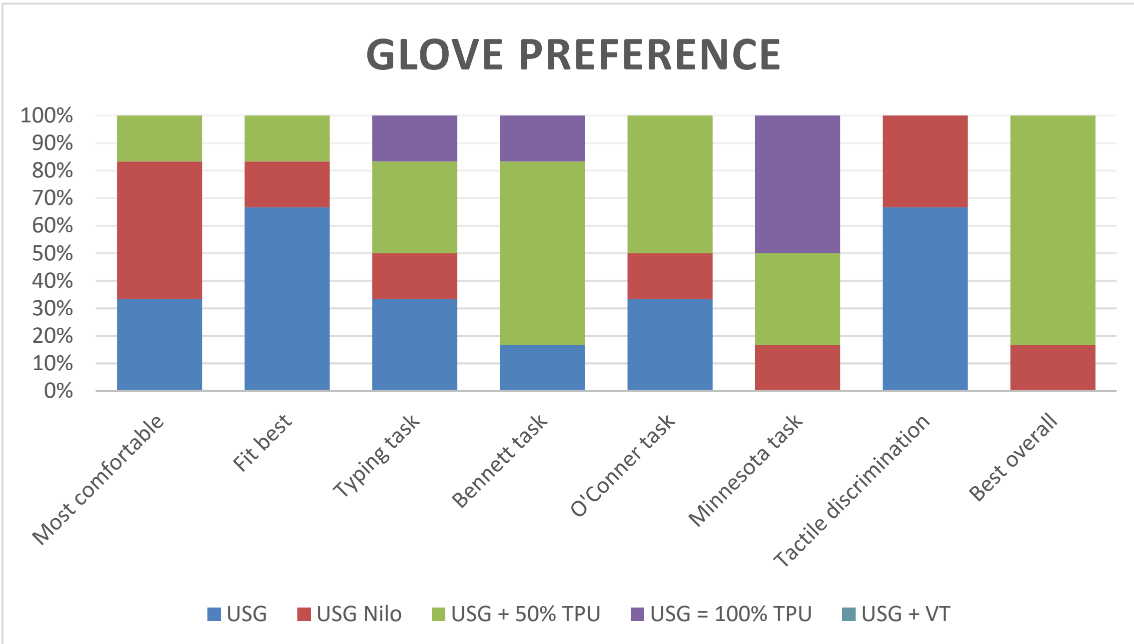


Figure 103 – Aggregated answers for post-session interview questions about preferred glove choice for comfort, fit, and task completion.

Following the completion of the final tasks of the study, participants were asked to indicate which of the gloves they preferred for fit, comfort, task completion, and as their overall choice for the entire study. Participants selected 4 different gloves as the dominant answer across all eight categories, though there were 2 clear winners.

The USG Nilo glove was selected for "most comfortable", confirming many of the comments made by participants in favor of this variant over the course of the study. The USG glove won out for fit and the Tactile discrimination test, reinforcing the benefits of the conforming material and sensory transparency that this design offers. The typing task question showed participants had no clear preference, as every glove received a vote except for the USG + VT variant.

The remaining task completion preferences all went to the TPU laminated variants, as these gloves provided a superior grip for tactile manipulation. The 100% TPU variant was preferred for the Minnesota test, but the 50% TPU variant was preferred as the majority favorite in the remaining tasks, and as the overall winner.

Table 10 – Aggregated values for post-session interview questions, with totals. Values shown indicate the number of participants who picked each glove as their preferred option for each question about comfort, fit, and task completion.

Preference	USG	USG Nilo	USG + 50% TPU	USG + 100% TPU	USG + VT
Most comfortable	2	3	1	0	0
Fit best	4	1	1	0	0
Typing task	2	1	2	1	0
Bennett task	1	0	4	1	0
O'Conner task	2	1	3	0	0
Minnesota task	0	1	2	3	0
Tactile discrimination	4	2	0	0	0
Best overall	0	1	5	0	0
Total	15	10	18	5	0

The USG + 50% TPU and USG Nilo gloves each received preference votes in seven of the wight questions, while the USG base glove received 6. This result for Study 3 showed a wider dispersal of votes than the results for Study 2 did – which showed a strong preference for the USG/Powermesh glove across almost all categories. It is notable then that the USG

+ 50% TPU variant scored more votes in Study 3, as the USG glove, tasks and participants were all identical between the studies. This does seem to indicate a clear preference for the balance of grip and sensation offered by the TPU glove, when the original preferred glove was available as an option.

7.5.2 Glove observations

One notable result of this study is the consistent appreciation and performance of the USG Nilo glove. The Nilo variant outperformed the USG glove on almost all task metrics and received preference votes across almost all task categories. These results can be contrasted with the performance of the Running glove in Study, which the Nilo glove approximated. The Running glove rarely outperformed the USG/Powermesh glove in Study 2 and received only 3 total preference votes in the final tally, compared to the 10 received by the Nilo glove. Participants noticed these similarities and explained that they appreciated the soft finish of the Nilo fabric, with some mentioning that they preferred the feel of it over the Powermesh. Participants also indicated that they preferred the USG Nilo gloves to the running gloves as they didn’t have the conductive fabric patches over the thumb and index fingers, as the Running glove did, and found the Nilo fabric to have better texture and grip, while still being thin enough to feel through. In both cases, the USG Nilo gloves were able to offer better grip than the Running gloves, while still maintaining improved levels of comfort and sensory transparency.

This contrasts with participant’s views on the USG gloves, which while comfortable and transparent, were seen to be the worst performers for feeling slippery during object interactions. This slippery finish is mentioned by all participants, each across most of the tasks, and seems to dominate the discussion of these gloves, beyond the topics of comfort, breathability, and transparency.

The TPU gloves are seen as the solution for the slippery surface issue, but they come with their own challenges. The USG + 100% TPU variant were praised for their added grip, but this came at the expense of finger movement. Participants reported that the TPU inhibited the natural stretch of the Powermesh, and this hindered their freedom to move their hands. The TPU was also reported as keeping the glove fingers stretched out, which left the participant's fingers suspended away from the end of the glove fingertips. Participants also reported that the TPU bunched in their palms and got in their way while performing gross grasp movements. Finally, the TPU was seen to inhibit their ability to sense shapes, as it made it harder to detect edges and corners. These observations mostly stem from the added stiffness and thickness of the material, which noticeably altered the natural characteristics of the Powermesh substrate.

The USG + 50% TPU glove was designed to strike a balance between the USG glove and the USG + 100% TPU glove designs, intending to offer participants added grip, while preserving the sensory transparency of the Powermesh. This appears to have been a largely successful effort, as the 50% TPU variant performed better than the USG and USG Nilo gloves in the Bennet and Minnesota dexterity tests – where the added grip provided noticeable benefit – and comparable to their results in the Typing and O'Conner tasks that benefitted from enhanced sensing. The consistent preference votes also show an appreciation for this combination of features, along with the highest preference vote score, and the selection as the overall preferred glove by 5 of 6 participants. This seems to indicate that by quantitative and qualitative measures that the USG + 50% TPU gloves offered the best overall set of features to their users.

By contrast, the USG + VT gloves scored the lowest on almost every metric they were measured in and were the only gloves to receive zero preference votes. The VTs were seen by participants as unnecessary burdens – weighing their fingers down and hiding the details of objects in their grasp. It is important to recognize that the VT gloves offered no haptic benefits in exchange for their encumbrances, so this is a challenging set of conditions for these gloves to be tested under. However, it is clear the VTs offered no real advantage to physical object interactions for the participants. The primary adaptation that was mentioned to allow them to deal with the VTs was to attempt to move them out of the way of their fingertips, or to use parts of their finger that weren't covered by the motor. One participant summed up their feeling about the VT gloves by saying "It almost feel like these are worse than the hockey gloves".

7.5.3 Itchy hands

For the first time in this series of studies, participants talked routinely about feeling that the gloves were making their hands feel itchy. Three of the six participants were particularly aware of this phenomenon, and all attributed it to being able to feel the seams of the gloves rubbing against their hands. The wrist seam was singled out as a location where these participants noticed the itchy feeling the most, though the thumb and fingers each received their own mentions. These observations applied to all gloves in the study, which is notable as they were all built on the same base USG pattern. It is possible that this sensitivity is due in part to the nature of some of the tasks, which asked participants to focus on very small stimuli, like the small pins and shape edges.

7.5.4 Conclusion

Study 3 was structured to test the ability of the Cost of Haptics protocol to evaluate gloves with smaller differences between them than the original Commercial glove set, and to test whether the protocol was capable of evaluating gloves that varied in their designs by a single independent variable. In this regard, the study was successful, as it returned results showing differences in performance and perception between these otherwise similar gloves. This was a positive result, especially given the added challenges of running the remote protocol, and the small N of 6 that the study was required to run with. As discussed in Chapter 6.5.7, this small sample was the consequence of conducting this research during a pandemic. A larger, more diverse sample may provide results that are more indicative of the experiences of the broader population.

This study showed the first example of testing a design variant that had been built on top of findings from previous results – here proving out the positive performance of the USG + 50% TPU glove. The underlying effects that informed the creation of this glove were first observed in the earliest pilots for the original studies, so it is gratifying to see a design hypothesis successfully tested, and to see it perform as expected. This shows the potential for this protocol in testing iterative design variants, and shows that it is sensitive enough to explore the differences in performance between even broadly similar gloves.

This study does also show the resources required to complete research focused on custom gloves. Testing five design variants at five different sizes, with tasks that need two hands to complete, meant that fifty gloves were necessary to begin to run it. As the remote study logistics required a contingency plan, this meant that 100 gloves were ultimately needed. Studies of this nature are therefore somewhat limited to teams who have the resources to build this expansive set of gloves, and to iterate further on their findings.

CHAPTER 8. DISCUSSION AND FUTURE WORK

8.1 The Cost of Haptics

Through the three studies discussed in this project, it is clear that the users of these gloves have born various costs imposed by the gloves – particularly seen as a wide variety of encumbrance effects. These costs are understood as tradeoffs for the benefits that the gloves provide, and the haptic capabilities of VR gloves are likely to provide a great deal of benefit to their wearers. The challenges with dexterous object interaction that were observed during the studies are unlikely to have much negative impact on VR glove users, as tangible object interaction is not the focus of most contemporary VR systems. Yet adapting haptic gloves to MR contexts introduces object interaction as a primary focus, which means the effects observed during these studies are likely to be observed in MR glove interactions as well. So, what are the costs of haptics for MR gloves?

The gloves tested in these studies contained three exemplars of haptic gloves. The hockey gloves carried their bulk on the back of the hands and fingers in a manner similar to the designs of contemporary pneumatic gloves, such as those from HaptX. The tactical gloves represent a slimmed down haptic glove design, that might be used in conjunction with thin-film haptic transducers. The USG + VT gloves represent a basic class of glove that places vibrotactors directly in place over the fingertips – a design common to lower-cost gloves, and the design used for the author's first design explorations. To have a sense of the likely encumbrance costs imposed by these types of MR haptic gloves on their wearers, we can therefore examine the performance of these three gloves across the studies.

It is also useful to contrast these three gloves with the highest-performing and most preferred gloves – notably the User Study Glove (USG) and the USG + 50% TPU variant. Participants confirmed, through their performance and preference that these gloves were far less encumbering than the haptic exemplar gloves and would be better suited for MR interaction contexts.

8.1.1 Sensation vs. Manipulation

The encumbrance effects imposed by the material finish of the gloves during object interactions was one of the dominant themes across all three studies. Participants freely reported the slipping sensation they felt wearing the gloves as their tacky skin surface was covered by smooth or blocky glove features. Consistently, the smoother slipperier gloves were also the set with the thinnest fabrics, and the greatest sensory transparency. The gloves with grippier surfaces, like the Gardening and Tactical gloves, tended to also have thicker material over the fingertip.

In Study 2, this created a trade-off in the commercial gloves – the Gardening glove was preferred for dextrous manipulation tasks, and the USG was preferred for the remaining sensing-oriented tasks. Object interaction through the tests seemed to require both sensing and dexterity for consistent control, which suggests that designers of MR gloves will have to weigh the benefits of designs features tailored to each capability. Finding a balance in the glove design between thin material over the fingertip, while still providing grip augmentation is likely the strategy that will lead to the most consistent object interaction for the wearer.

Study 3 showed one possible solution with the introduction of the USG + 50% TPU variant. By splitting the surface area of the glove 50/50 between the thin Powermesh material and the grippy TPU, participants were able to achieve consistent performance in both the dextrous manipulation tasks and sensing tasks. Participants ultimately selected this design as their overall preferred option. This is notable as the basic User Study Glove was the unanimous overall favourite in Study 2 and was tested again in Study 3 with the same tasks, questions, and participants. The only difference between the Study 2 winner and the Study 3 winner was the addition of the perforated TPU laminate, which suggests a strong recognition from the participants of the advantages of this specific design feature.

The haptic exemplar gloves did not fare as well in these comparisons. Both the Hockey and USG + VT gloves were considered to be slippery due to their bulk preventing the participants from maintaining a firm grip on the test objects. The VT was prone to moving in the pocket it was embedded in, which added to the unpredictability of the grasp. One participant offered this retort as they struggled with the USG + VT gloves: "Why on earth would you take the most important part of your finger and not let people use it?". With these two gloves, the first cost of haptics was the participant's inability to maintain effective control of objects while wearing them.

8.1.2 *Errors and variability*

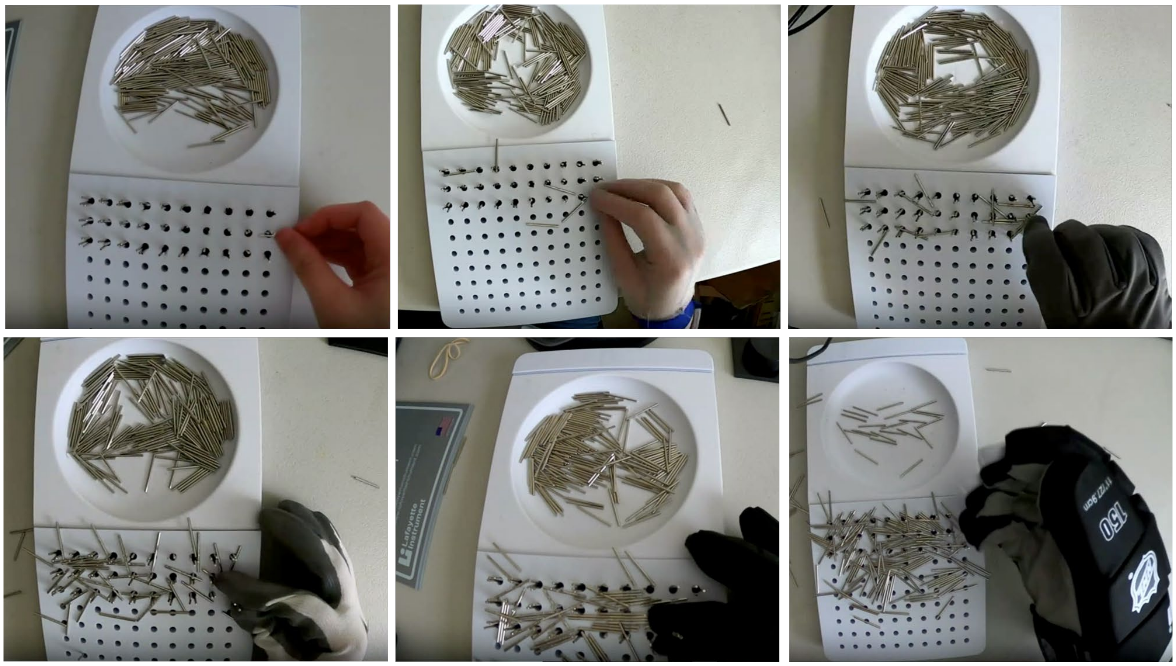


Figure 104 – Screen captures of the O'Conner dexterity test, showing the count of dropped pins across the same participant. There is an obvious increase in the count of dropped pins as the relative bulk of the gloves increases.

One byproduct of the decreased consistency in object manipulation and control was the increases in error rates, and performance variability – both seemingly obvious signs of encumbrance. The O'Conner dexterity test, for example, had a median error rate of < 10 dropped pins for the "No Glove" condition, and a median rate of ~70 dropped pins for the Hockey glove. This seven fold increase was due to the participant's inability to keep the pins from slipping out of their fingers, and the increased difficulty in collecting them once they had fallen. The dramatic difference in dropped pins for each glove is shown in Figure 104.

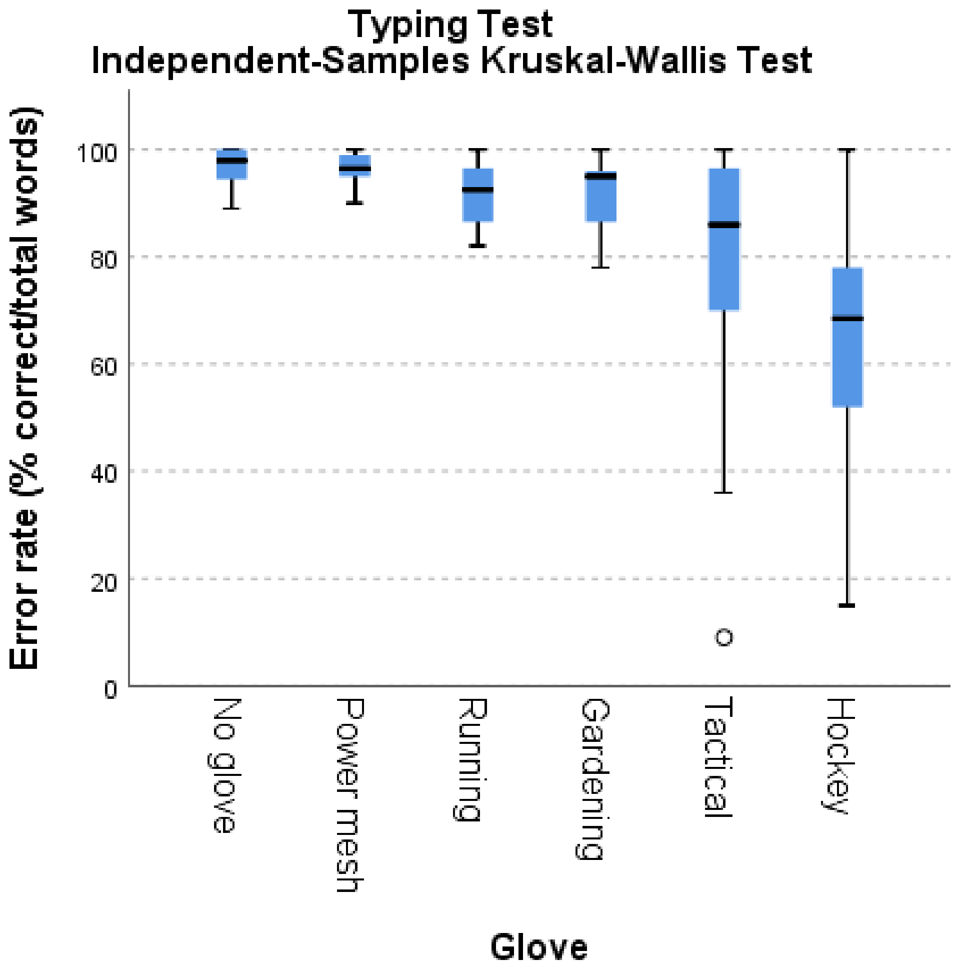


Figure 105 – Typing test error rate Kruskal-Wallis test results – there is a significant difference in the error rate between gloves, $\chi2(5) = 27.12$, $p = 0.000$

An additional example is found in the typing test – the error rate with the Hockey gloves stretched from ~15% correct words to 100% correct. This contrasts with the "No Glove" condition, which ranged from ~90% to 100% correct words, with the median close to 98%. The variability of the Hockey glove result is likely a product of the glove's impairment of the participant's normal typing skills, and their hurried adaptation to find a workable replacement strategy.

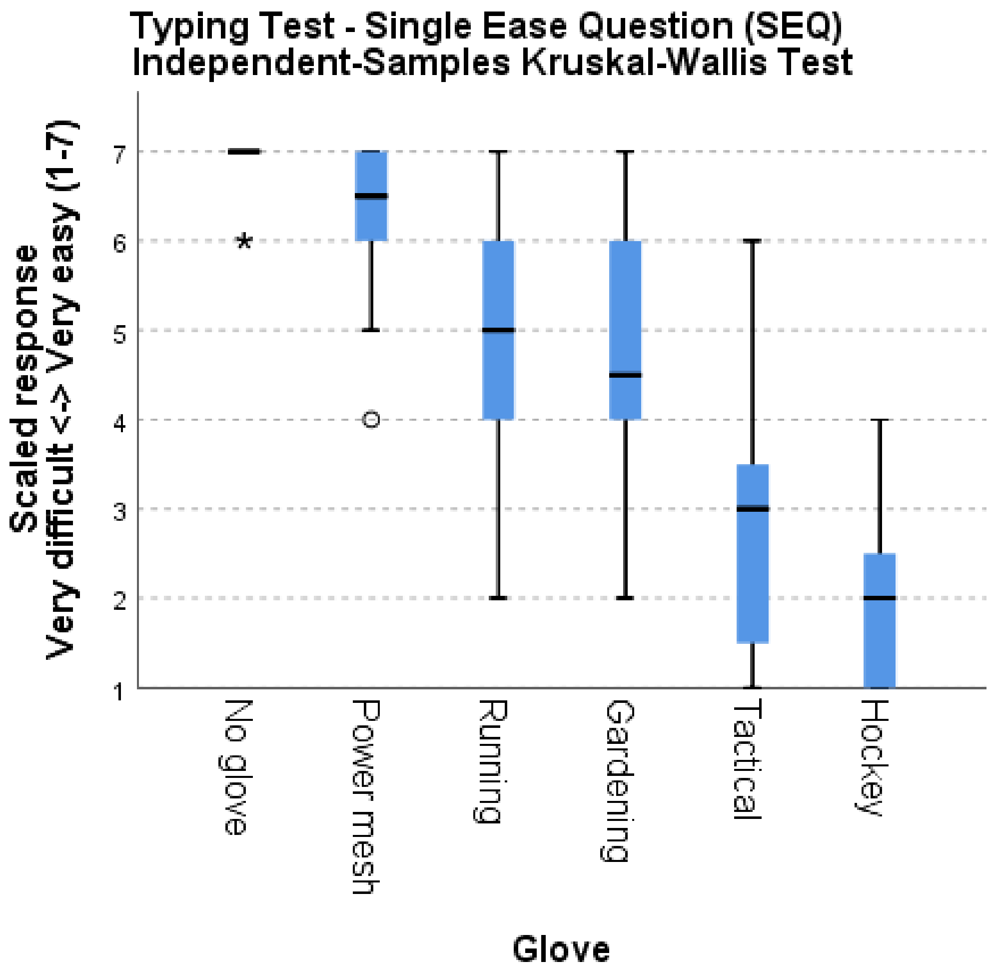


Figure 106 – Typing test Single Ease Question Kruskal-Wallis test results – there is a significant difference in the perceived difficulty of the task between gloves, $\chi2(5) = 15.95$, $p = 0.007$

Completing tasks with high-encumbrance gloves makes the task experience much more unpredictable for the user, as they seek to overcome the impairment of their normal dexterous manipulation capability. This naturally leads to a corresponding perceived increase in the difficulty of the task, shown in Figure 106, and frustration for the user. The second cost of haptics is therefore the unpredictable impairments of the wearer's normal skills, and an increase in the number of errors they make.

8.1.3 Fingernails

Figure 107 – Photo examples of participants using their nails to test for specific physical features in the small scales of the Tactile Discrimination test.

Many participants discussed wanting to use their nails through the gloves to help get better leverage on an object they were trying to pick up – something that was possible with the stretchy knit gloves, and more difficult with the heavier gloves. The Tactile Discrimination test was an example of this – particularly the hollow square condition – as participants used their nails to drag over the surface of the feature – as illustrated in Figure 107. Feeling their nail snag on the edge of the concave feature seemed to amplify their ability to sense it, and most participants adapted to start using their nails periodically throughout the task.

Figure 108 – Photo of a participant's hand, showing the length of their enhanced nails.

Participants with long nails, as seen in Figure 108, also struggled with some tasks as the gloves created a bubble of air on top of the underside of their nail. This made it harder to grasp objects, and harder to identify where the control surface of the finger was under the glove. This hollow void created inside the shell of the nail presents a particular problem for smart glove designers, as the sensors and actuators intended to interact with the wearer's fingertips will be situated over this void when worn over long fingernails. This is disadvantageous to any groups who typically wear their nails long, and will need to be closely considered by a glove design team.

The third cost of haptics is therefore the loss of enhanced sensing and manipulation by the user's fingernails, and the interaction component misalignment for users with long nails.

8.1.4 Fit and comfort

The participant's fit and comfort were frequently discussed over the course of the studies. Most commented with clear distinctions between the light-weight stretchy gloves, such as the Powermesh gloves, running gloves, and Nilo gloves, with the heavier Tactical and Hockey gloves. The latter set scored poorly in these assessments across the board, likely due to their bulk, stiff materials, and restrictions on movement. The former group were all constructed from extensible fabrics – capable of flexing and conforming to the wearer's hand.

The most common complaints about comfort were linked to thermal management and seams. Most participants commented on their hands getting hot and/or sweaty while in session in the studies, due to their exertion in completing the tasks, and the length of time they were wearing the gloves. Sweat is an issue that needs to be considered by smart glove design teams, as the build-up of moisture in the glove may interfere with electrical components, and the trapped moisture creates discomfort and possible issues with hygiene.

Participants also discussed the seams of the gloves at length, repeatedly calling out the more prominent seams as "itchy" and "bothersome". All tested gloves had noticeable seams to various degrees, with the exception of the Gardening gloves, which were manufactured using a seamless knit process. Seams appeared to interfere with users perceptions during delicate sensing tasks, and they led to mild discomfort. Multiple participants in Studies 2 and 3 referred to feeling itchy and irritated after an extended study session.

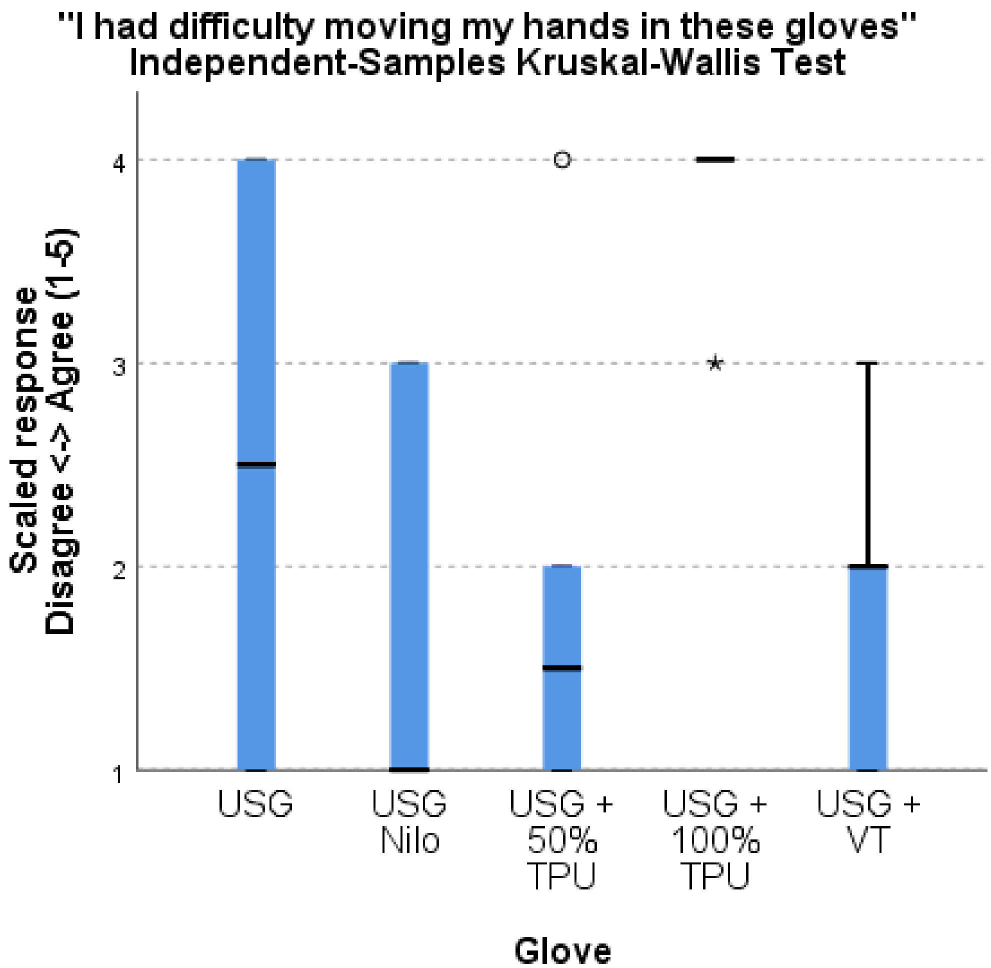


Figure 109 – Scaled response for “I had difficulty moving my hands in these gloves" Kruskal-Wallis test results – there is a significant difference in agreement with the statement between gloves, $\chi2(4) = 11.94$, $p = 0.018$

Fit issues manifested in two general categories – overly tight fits that led to motion restriction, and overly loose fits that led to bunched fabric that got in the participant’s way. The lighter knit gloves were again preferred in this area, as their stretchy fabric was more forgiving with tight fits, and inhibited movement less. A notable exception to this is the USG + 100% TPU glove, which participants singled out for limiting their movement, as seen in Figure 109. This result contrasts with positive perception of the USG + 50% TPU glove as the only variation between the two glove designs is the perforations in the TPU laminate.

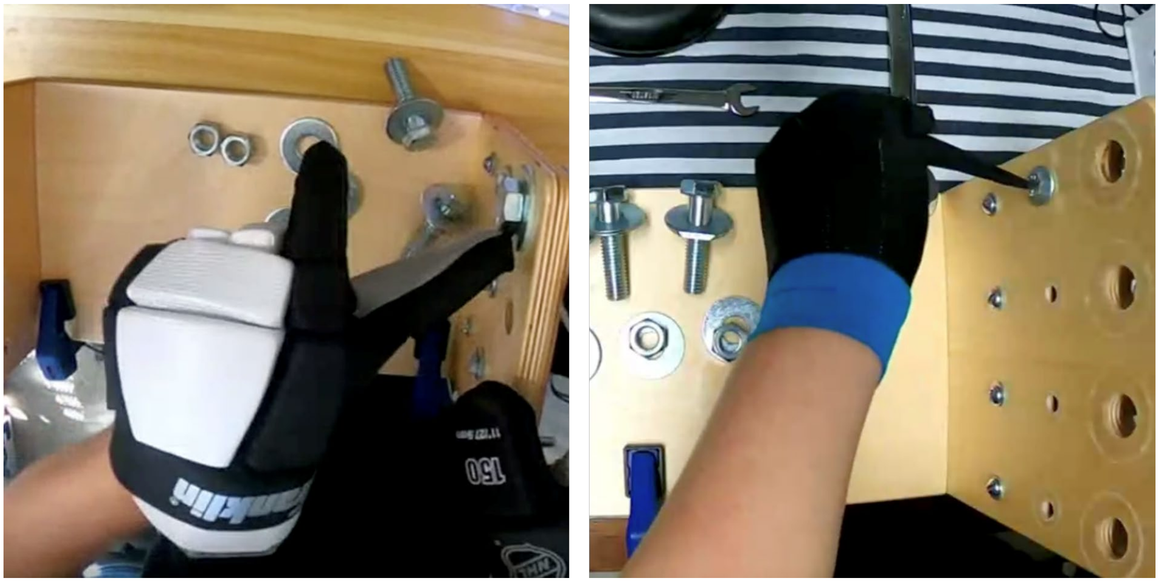


Figure 110 – Photos showing the Hockey and Nilo gloves being snagged on a bolt during the Bennett Dexterity test, due to excess fabric at the fingertips.

Participants also experienced adverse effects due to overly large fits – notably the bunching of excess fabric at their fingertips. This was observed to interfere with the stability of their grip on objects, but the excess fabric was also prone to snagging and getting caught on object edges – shown with photos of two glove fingers being caught on the Bennet test in Figure 110. Participants self-selected the size of glove out of the available options, yet reported that they were frequently unable to find a size that fit them well overall – with excess finger length as one of the most common examples they cited.

Fit and comfort therefore provide an additional three haptic costs – increased heat and sweat for the wearer, decreased finger mobility, and excess fabric bunching and snagging on the environment.

8.1.5 Adaptation

Participants were observed to respond to the encumbrance effects imposed by the gloves by employing a wide variety of adaptation behaviours. Some of the simplest responses were focused on adjusting the gloves to fit – pulling the cuff or fingers down to try and reduce excess fabric, or to relieve a tight spot that had developed. Other participants focused on changing their grip – with the numerous grasp interactions in the Box and Blocks task, participants reported trying out various grasp strategies, as they varied the number and orientation of the fingers they were using. Another reported strategy was to grasp harder to overcome slip – a common explanation of their response to the Maze Tracing task.

These adaptations may be expected, but they were observed far more frequently with the glove conditions than with the “no glove” condition. Participants frequently found it difficult to describe their strategy while performing their tasks with bare hands – a common explanation was that they didn’t need to think about the task much, as their body knew what it was doing. In contrast, while experiencing increased encumbrance effects with the gloves, their natural behaviour was interrupted, and they needed to come up with a new strategy during the operation of the task. Participant’s reported needing to think about tasks more than normal and described some degree of additional cognitive load and frustration.

One clear example of this adaptation burden could be found with the with the Hockey gloves in the Typing test. As participant's natural touch typing ability was affected by the gloves, they started to adapt their typing style to a slower "hunt and peck" strategy, which dragged their view down to the keyboard, and away from the screen. However, as the bulk of the Hockey gloves obscured their view of the keyboard, they could not rely on pressing the key to confirm they had correctly typed the right character, and needed to return their gaze to the screen. This greatly increased the worked required by the participant for every single key press and explains their high Frustration score for the task. This is therefore an example of the visual mass of the gloves providing their own encumbrance to the user, as they occlude and inhibit visual confirmation of the object in their hand.

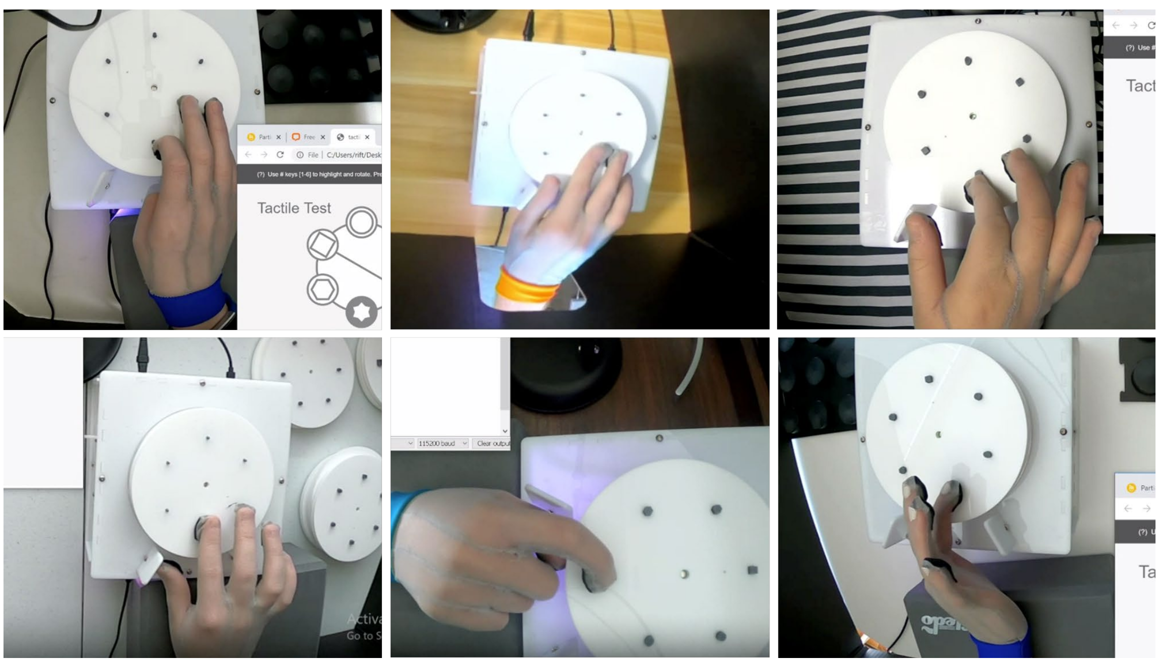


Figure 111 – Screen captures showing video still of participants adjusting the VT glove to move the VTs out of the way during the Tactile Discrimination test.

An additional type of adaptation was observed during the Tactile Discrimination test, while participants were wearing the USG + VT gloves. The VT completely obstructed their tactile sensing capabilities, but it was attached to the glove fingers by means of a small sewn pocket. This pocket in combination with the underlying stretchy mesh fabric, allowed the participants to manoeuvre the VT away from their fingertips temporarily, to better sense the test shapes they were trying to identify. The participant's ability to shift components out of their way allowed them to better complete the task and reduced the otherwise debilitating encumbrance effects imposed by the glove design. Given the total sensory occlusion of the VT, this flexible adjustment strategy allows the wearer to bypass design features that could lead to task failure.

8.1.6 Costs and Trade-offs

With consideration of the findings from all three studies, a final aggregated list of the costs that haptic gloves may impose on their wearers is as follows:

1. A loss of consistent tactile control over the physical objects they are manipulating with their hands
2. A loss of haptic sensation to observe and discriminate features of the objects they are interacting with
3. A perceived increase in difficulty with tasks they are otherwise capable of completing
4. An increase in the number of errors they commit while trying to complete an otherwise simple task

5. An increase in the variability of their task performance, lowering the predictability of the outcome of a common task
6. The occlusion of their fingernails as sensing and manipulating tools
7. If the wearer has long fingernails, the glove fit will be affected – interfering with the accurate placement of interactive components, and creating a useless void in the gap between their nail and the glove fabric
8. An increase in the thermal discomfort experienced by the wearer, and a corresponding increase in sweat captured by the gloves
9. An increase in the discomfort due to abrasion by seams and other components and construction elements placed on inside of the gloves
10. With an overly tight fit, there may be an increase in the glove's inhibition of movement by the wearer
11. With an overly loose fit, there may be an increase in the amount of fabric that bunches at the fingertips, which is likely to decrease normal sensation and control capabilities, and snag on objects and the environment
12. The impairment of the performance of normal tasks with which the user has developed a skill for (such as typing)
13. An increase in the cognitive load and frustration experienced by the user as they seek for new strategies to overcome the loss of their inherent skill
14. An increase in the need for visual confirmation of the state of the object they are manipulating, to offset the loss of haptic feedback from the object

15. An impairment of their ability to visually confirm the state of objects held in their hand, due to the bulk of the glove occluding the object

Individually, these costs may be manageable for the participants in these studies, but collectively they create a significant encumbrance on the user, which has a debilitating effect on their ability to complete common object interaction tasks. However, there are tradeoffs in play here, as these costs are associated with the user's completion of tasks focused around physical objects. These costs are borne by the user in exchange for the benefits provided to them by the enhanced capabilities of the gloves they are wearing. The two haptic glove exemplars – the Hockey and USG + VT gloves – clearly imposed more of these costs on their wearers, but the gloves they represent give them additional haptic capabilities than they would otherwise have. In VR, this may be an acceptable tradeoff, but in MR these costs will heavily bias their experiences towards virtual objects. To create a set of gloves that allows equal unencumbered access to both virtual and physical objects, the design team may want to consider the following recommendations:

8.1.7 Design recommendations for MR gloves

Gloves for MR should seek to limit the amount of sensory occlusion they impose on their wearers – either by avoiding placing material over key sensing areas, or by using lightweight, stretchy, and otherwise sensory transparent materials. This material selection will also benefit the thermal comfort of the wearer and should limit excess sweat buildup. If this is not possible within the design constraints, an alternative strategy built around moisture wicking should be investigated. Where possible, the gloves should limit or eliminate internal seams, and investigate using seamless continuous knitting fabrication strategies, to avoid excess abrasion on the skin.

The surface of the glove should be treated to add flexible material that enhances the grip and tactile control of objects by the person wearing the glove. This grippy material should not further inhibit the sensory transparency of the glove, so pattern strategies such as perforation or screen printing should be explored. Where possible, the design of the glove should feature a thin profile over the fingertips – to avoid excess bulk that may interfere with object manipulation, or occlude the view of the object while in the hand.

Gloves designs for MR should consider the adaptive needs of their wearers, who may need to employ various strategies to adjust the gloves to fit their current task or environment. Modular or adjustable solutions for interactive component placement should be considered, to allow the wearer to make quick adjustments or customizations that allow them to bypass encumbering features when necessary. Accurate fit should be prioritized, and parametric or custom fit strategies should be employed to ensure accurate placement of components. If standard sizes are employed, they should be offered in smaller graded steps than conventional gloves, to ensure that fit can be optimized. These variations on pattern grading should also consider wearers with long fingernails and offer standard sizes that feature longer and shorter fingers. Matching the length of the finger to the wearer, and placing components accurately for each individual should be a priority.

These recommendations should provide a starting point to create useful and usable gloves, but all design features should be tested to evaluate the encumbrance costs they impose.

Beyond physical impairments, care should be given to observe any effects the glove may have on the wearer's dexterous skills, fatigue, frustration, and cognitive load.

8.2 Future work

8.2.1 Research Prototype evaluation

Following the direction of the three studies that formed this document, there is one additional future study that could be planned. This is one focused on the evaluation and benchmarking of completed glove research prototypes. In contrast to the constructed glove study – which shows the application of this evaluation approach at the beginning of a design process – this study could focus on late-stage evaluation of complete design prototypes. This would allow for benchmarking of milestone releases and finished products, and could show the overall progress of the design program over time.

This study would share the "no glove" and "basic glove" conditions from the constructed glove study, but add two additional gloves – a "low encumbrance" model and a "high encumbrance" model. These might be drawn from projects in process with an existing glove design team, or historical exemplars. The "low encumbrance" model would demonstrate a light-weight form, with a focus on gestural capabilities, where the "high encumbrance" model would focus on VR believability, and offer a higher-fidelity haptic display. This would stake out two high-value spots on the smart glove encumbrance spectrum.

As with the constructed glove study, this study could test a series of possible hypotheses:

H_0: There will be no significant difference in performance between the glove conditions.

H_1: There will be a significant difference in performance between the glove conditions.

H_2: The “no glove” condition will have better performance across all tasks than any glove condition.

H_3: The “basic glove” will have better performance across all tasks than either smart glove.

H_4: The “low encumbrance” smart glove will have better performance across all tasks than the “high encumbrance” smart glove.

This study could follow the same protocol as the prior studies and would likely offer similar results. The key difference is the opportunity to evaluate completed glove designs, with the goal of making longitudinal comparisons. This would allow for the establishment of benchmark scores of completed designs – either to compare to iterative generations of the same design, or to form this basis for new product goals. This latter option would allow glove projects with performance improvement goals to be made – for example, designing a glove that improves on the median typing speed of a prior version by 10%. This creates the opportunity to structure more quantitative requirements, and to better track progress in the design revisions as they are completed. Continuously following this approach would quickly lead to the establishment of a benchmark library, which would allow for the comparison of new designs without the need to re-test older ones.

8.3 Encumbrance Evaluation Framework

The studies in the project ultimately showed there were significant differences that could be detected between the glove selected for the study, which allowed for detailed comparisons within the set. However, the findings from these studies are complex, as they were presented as a set of metrics based on results from the study results of dozens of different measures. While these metrics allow for direct comparison between designs, it is harder to aggregate them to get an overall sense of the encumbrance performance of each glove. In the future, there may be an opportunity here to synthesize these results, and develop a standardized scoring and benchmarking system. A system like this would better support direct comparison between designs and more easily provide historical comparisons between previous versions.

An example precedent of this approach may be the NASA-TLX scale, which individually measures six different characteristics of user load, in individual scales, but then offers a mechanism to provide a single weighted score that can be more directly compared [30]. NASA-TLX scores are not easily compared across unrelated studies, but within a set of studies comparing similar designs the tool can be very useful to judge relative performance. Employing this same model for glove encumbrance should allow for similar comparisons between gloves, using an easily accessible score.

The scale scores could be calculated – following the model from Godfried Augenbroe's design performance work – by establishing the performance indicators that most directly indicate each scale and normalizing them to fit within the same range [4]. These individual scale scores already will summarize a rich data set underneath them – formed from both the quantitative and qualitative results gathered through the study – but could also be aggregated into a single encumbrance score. One step that may ease this aggregation process into a final score is the establishment of a set of standardized scales.

8.3.1 *Performance scales*

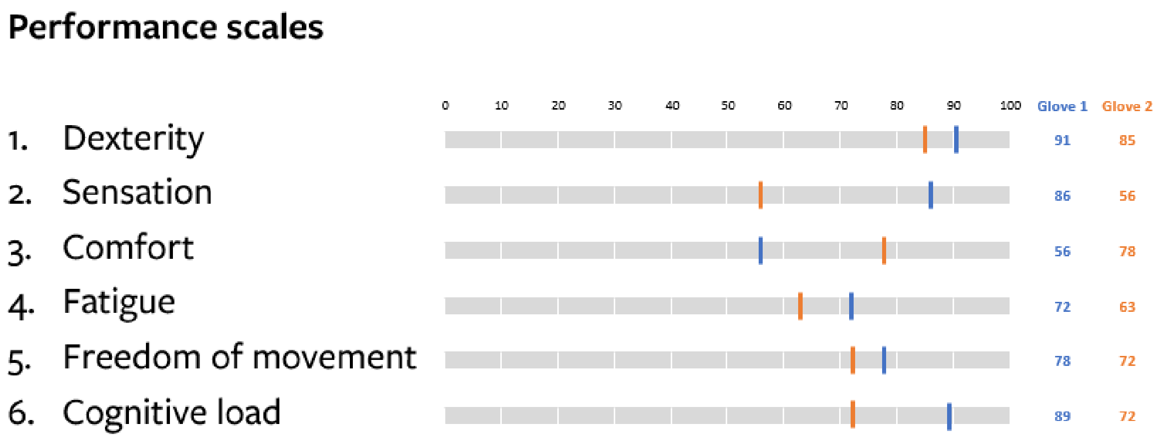


Figure 112 – An illustration showing the potential design of the primary six scales in the proposed framework

These scales would each be indicative of high-level clusters of performance metrics, aggregated into a single score. This would involve identifying which performance metrics are most useful to cluster together, and which ones show a natural affinity to each other. One hypothetical structure for these scales – a illustrative example shown in Figure 112 – could be drawn from the earlier cited literature review from Dianat at al [19]. These categories of tests were drawn from the existing literature of glove research and are indicative of broad clusters of investigation. The six scales indicated would seem to be broadly useful as aggregate scores, but the compiling of these scores may prove challenging.

The performance data from each metric would need to be normalized before aggregation. A possible method for determining the normalized values would be to set a range from 0 – 100, where 100 is the mean performance of the baseline condition for each scale, and where 0 is the worst expected performance for the most encumbering condition. This would constrain the scale to a useful range of encumbrance, with enough gradation to allow for detailed comparison. However, it would still be possible to have performances that exceed the bounds of the worst-expected condition, or to have results where the highest performing condition isn't the baseline, which would provide for results that fall outside of the normal range.

Other challenges that would need to be overcome are concerned with the relationships between the normalized results from each metric. If they do not consistently co-vary, they may not group effectively together, or even cancel each other out. It may also be necessary to weight the metrics disproportionately, as they may not contribute to the overall scale in the same ways. It is also not clear how to treat the mixed-methods results of these protocols, as structuring each scale would likely involve the aggregation of both nominal data (such as typing speed) and ordinal data (such as scaled response questions), which may lead to misleading results.

8.3.2 Performance charts

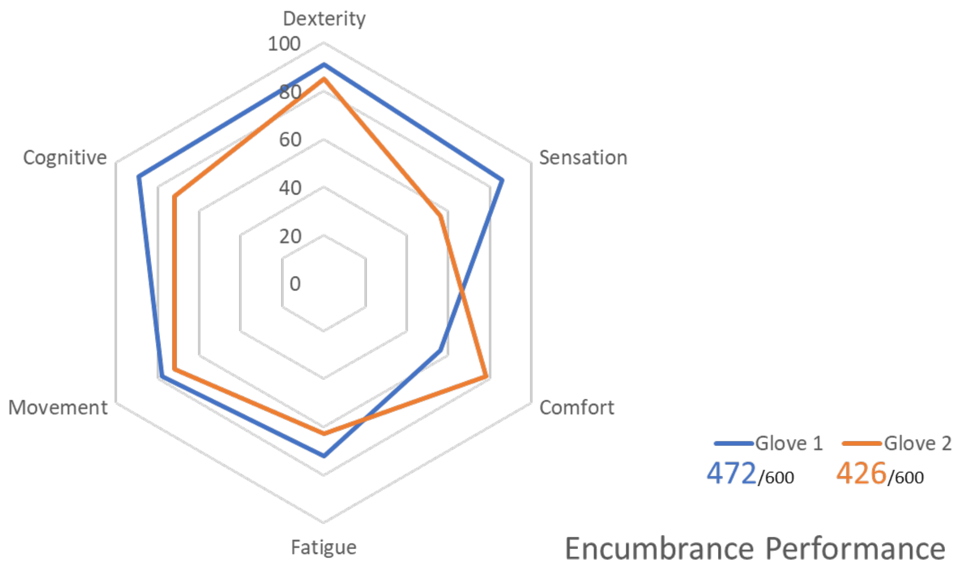


Figure 113 – A mock-up showing a potential method for rendering the encumbrance scores from this proposed study in a radar chart.

Following Augenbroe's approach, illustrated in Figure 113 – if the scales are established, they can be charted using a radar chart. This charting creates a visual record of the performance of each glove and allows for easy comparison of multiple gloves or design alternatives on a single chart. Radar charts are suitable in this application for showing multiple sets of multi-variate data on top of each other and can show outliers and similarities in performance between two gloves. This provides many benefits for presenting glanceable data and seems well suited to the target audience of designers. However, more in-depth analysis should be done with the raw data, as comparing the areas of two conditions in the radar chart can be misleading.

8.3.3 Performance score

A unified performance score would be the final proposed feature of the analysis system – a single number that can communicate the relative overall encumbrance performance between two gloves. This score can be calculated based on the scales, or the underlying data beneath the scales, but it should allow for the easiest comparison between two design alternatives, and is the most basic representation of the encumbrance cost of the design.

The score could provide the most easily understood basis for any benchmarking between gloves, and for longitudinal comparisons. Constructed gloves, glove prototypes, and historical benchmarks could be reliably compared, and the score could infer direct performance results – for example, a score that is twice that of a competing design should offer twice the performance.

8.3.4 Framework evaluation

It is clear that there would be a great benefit in communicating and comparing the results of these studies if the simplified scale and score structure was put into place in the framework. This system would also greatly benefit the target population of designers, many of whom are not well trained in statistical analysis. However, it is also clear that creating a reliable and meaningful set of scales would involve substantial development and testing. Calculating this score would also involve many of the same risks and complexities as the initial calculation of the underlying scales.

The scale scores are likely best understood as ordinal numbers, which means they cannot be simply averaged to provide a meaningful result. To arrive at a single score is to take a few dozen independent measures across a disparate range of effects, and attempt to gang them together through a series of progressive aggregation calculations. After repeated normalizing, weighting, and averaging processes, it is not clear exactly what the final score would be representing, or whether it would be accurate and repeatable. This process might produce a simple outcome, but it is likely to be a misleading one, without extensive continued research.

However, these challenges do not diminish the opportunity to continue work on this project. The value of the simplified score is undeniable to the potential end users, and it may be possible to develop an app framework that automates much of the complex calculation to produce it. Tools like these would be invaluable in encouraging designers to run more small-N studies and basing more of their design selections on evidence.

8.4 Conclusion

The studies described in this document were created with the goal of identifying a test regime that was capable of measuring encumbrance effects of gloves. These tests should be capable of detecting significant differences in performance between gloves and be used to effectively evaluate and compare different glove designs. It is clear that the Cost of Haptics protocol was ultimately capable of meeting this objective, as the performance results for the gloves were routinely significantly different from each other across a range of quantitative and qualitative measures. These differences pointed to measurable impacts of various design features of the gloves and were shown to be repeatable across pilots and three successive studies.

Study 1 successfully replicated the results from the precedent 2018 pilot, which helped validate that the results were repeatable. Study 2 revised the protocol to deal with issues of external validity, and again showed results that were broadly similar to the findings from Study 1, even with a change in tests. Study 3 showed similarly significant results with a set of gloves that were far more similar to each other than the commercial set in Study 2. In all cases, there were consistent signals captured that indicated that a consistent set of encumbrance effects were being experienced by the participants.

It is notable that the protocol and measures selected were capable of showing significant differences between the similar gloves of Study 3, even with a modest N=6 participant count. While increased sample count would surely lead to improved reliability of the results, it is important to note that a small-N study is viable with this protocol, as that is a useful scale for small design teams to be working with. One of the challenges of using a similar protocol in an iterative design process is the need to test after every major iteration. The ability to build those tests around small-N studies like this one lowers the bar to entry greatly and should encourage design teams to test more frequently.

These studies also showed the value of a mixed-methods approach to this type of research. The quantitative results provided a reliable way to measure the size of the effect, while the qualitative results offered possible explanations for their cause. This allowed opportunities for triangulation and comparison – to show where performance and perception align, and where they differ. Qualitative inquiry also allowed for summative questions of preference, which were useful in determining the participant's holistic impressions of the glove experience. Additionally, open-ended discussion with participants frequently pointed towards new explanations for effects, and new opportunities for further research.

The second main goal for these studies was to identify the various costs of haptic glove features when interacting with physical objects, and to determine opportunities to improve designs for haptic gloves for MR. These studies were ultimately successful in this effort, and led to the identification of 15 different encumbrance costs borne by the user while wearing haptic gloves (Chapter 8.1.6), which in turn produced a set of design recommendation for MR gloves (Chapter 8.1.7).

These costs might have been determined through other observational methods, but it is important to note that there is underlying statistical data beneath each of them – indicating the significance of the difference between the gloves, and the relative strength of the effect. Having access to this data allows for more precise evaluation of specific design features, allows for quantifiable tracking of performance improvements over time, and promotes discussion in concrete terms of the impacts of the glove designs on their wearers.

One final outcome from the study was the development of a design proposal to address some of the measured encumbrance effects, and the successful testing of that design to show performance improvements. The User Study Glove + 50% TPU variant in Study 3 was first identified as a possible solution in the 2018 pilot, as it was clear how pervasive the slip/grip effects were to the participants of that study. A glove that balanced the often-contrasting needs of sensation vs. manipulation design features was worth testing, and it ultimately proved to be successful. The glove performed well across tasks and was selected as the overall preferred choice for the protocol by 5 of 6 participants. This demonstrates how data can be used to structure and test a glove design that varies by a single feature, and then be used to evaluate the design hypothesis embedded in that feature.

Encumbrance will continue to be a high-value research topic as wearables are developed to deliver ever more complex interactions to their wearers. The studies performed in pursuit of this dissertation show that a mixed-methods protocol can be assembled and tested to show these differences, identify the cost to the user of their haptic gloves, and propose design solutions to address these challenges. Those design solutions should lead to better experiences for the users of Mixed Reality gloves, and ultimately find balance in addressing the Costs of Haptics.

APPENDIX A. DESIGN EXPLORATION

The "Haptic Mirror Therapy Glove" project was run by the author as a graduate research study at Georgia Tech from 2012-2017. The project – as initially conceived – aimed to develop a smart glove that would assist patients with home therapy; yet it became a case study on the pervasive effects of encumbrance in the various design features. These initial attempts to build functional smart gloves informed later work to more rigorously quantify the encumbrance effects that were observed.

A.1 Haptic Mirror Therapy Glove

Recovery from a stroke poses a significant challenge for patients in treatment, both due to the difficulty of the rehabilitation process and the effects of the emotional and social strain that can follow the injury. Patients who suffer from a stroke experience significant change in their lives, and struggle to accept the differences that may now dominate their lives as they begin their recovery. Rehabilitation promises at least a partial return to their prior way of life, but it requires dedication and a significant investment of time to a series of continuously repeated tasks, designed to prompt the brain to remap the lost function. Sticking with a therapeutic regime offers the stroke survivor the best chance for rehabilitation yet doing so can be a monotonous task. Patients have a better recovery experience when these tasks are constantly changing, and when they are challenged in new and different ways, which stimulates the brain to learn more quickly and efficiently.

One of the most common effects of stroke is hemiparesis, or the one-sided weakness of the body, particularly in the upper limbs. 80% or more of stroke survivors experience hemiparesis to some degree, with 55% to 75% having some ongoing limitations [90]. Various training and stimulation methods have been developed to aid with the rehabilitation of a paretic limb, yet most are labor intensive and require supervision from a therapist, which can limit their availability. Given that the chance for recovery increases with the timeliness and intensity of the rehabilitation program, there is an increased interest in the development of therapeutic systems that could be used at home as way of maximizing the patient's potential for recovery.

Mirror therapy has been proposed as a potential solution, as it is cost-effective and portable. Clinical trials have indicated that it is a useful therapeutic intervention, demonstrating improvements in the range of motion, speed and accuracy of movement, squeeze strength, and improvements in motor function in chronic stroke patients [44]. Haptic feedback also shows new promise in therapeutic programs, aiding in motor learning functions through the directed application of vibration stimulus to the affected limbs. Clinical applications of both mirror therapy and haptic stimulus build on the demonstrated potential of the brain to respond to mirrored stimulus – regenerating lost neural pathways that weaken an affected limb by re-mapping stimulus presented by an unaffected limb. There is now an opportunity to investigate the incorporation of both methods into a single unified therapy, suitable for use in the home.

A.1.1 Objectives

The project was structured with the goal of building a haptic glove device that, used in conjunction with established mirror therapy protocols, could speed up the time to recovery and motor sensation in a limb weakened by the affects of stroke. This would be done by developing an interactive wearable device that becomes part of a home rehab program that complements the work being done with the user's therapist. This device would be part of a system that would assist and guide this process by providing the user with the opportunity to develop their own rehabilitation exercises based around objects found in the home, using their unaffected hand to guide their affected one. This system was intended to improve the rehabilitation experience by encouraging the user to perform their exercises more often, with an increased awareness of their progress, and increased agency and self-sufficiency in their development.

To accomplish this objective, an electronic glove needed to be designed, fabricated, and tested to perform the necessary interactive functions reliably and safely, and be built in a suitably durable manner to survive the clinical trial process. Design criteria would be established by seeking input from stakeholders, including therapists, stroke survivors, and clinical researchers as well as various technical subject matter experts. Wearable technology is an emerging field, and there were not established best practices for the creation of medical devices built with conductive thread circuitry and other wearable components. These would need to be developed, and appropriate technology identified and selected. Corresponding materials research would also be required, with an investigation into current rapid prototyping and parametric small-run production practices.

The original goal of the project was to demonstrate the efficacy of haptic mirror therapy glove, which would require a clinical trial. To test the performance of the glove in a home setting, it must first pass an efficacy trial in a clinical setting, and appropriate usability and functional requirements in the lab. Home use also means the glove would have to prove to be usable by patients with a paretic limb without assistance, and have its function and interface be easily understood. It would be necessary for the glove to attain a high level of aesthetic acceptance amongst patients and therapists and prove to be comfortable after extended use.

A.1.2 Design Requirements

These design requirements were established to guide the project:

- Self-directed rehabilitation: The user must be able to operate the device with limited instruction and assistance

- Affordance: The device must effectively communicate its function

- Exercise creativity: The device must assist the user in designing and performing therapeutic exercises in the home

- Adaptive assistance: the device must provide some measure for adjusting to meet the user's requirements as their condition changes

- User comfort: short and long term, breathable

- Aesthetics and semiotics: the form and ‘look’ of the glove should be acceptable to users for use in their home

- Construction: Glove sewn using lightweight sewn circuits

- Durability: The glove should be robust and protect the electronics and sensors.

- Easy to don and doff by users without assistance, on both affected and unaffected hands

- Accommodate wide variety of people: either adjustable or multiple sizes or a combination.

- Motors and pressure sensors must be accurately placed on the fingers to maximize system response and lower error rate

- Pressure sensors must be able to measure a range of sensitivity from a light touch to a tight grip

- Motors should be easily tunable to customize appropriate stimulation

- Easy to manufacture and assemble, compatible with current wearable products rapid prototyping processes

- Wireless: The device must be battery powered and communicate wirelessly

- Safety: The device must present no known health hazard to the wearer

A.2 Medical Background

It is not immediately obvious why the observation of a healthy limb, or the application of mild vibration to the skin of an affected limb would be useful in a therapeutic context without understanding neuroplasticity – the ability of the brain to change its structure and function, and to regrow lost neural pathways. While still an emerging field of research, neuroplasticity has upended the centuries old construct of the brain as series of fixed functional areas, hardwired like a machine, and where damage is permanent and irreversible [21]. Instead, we may be able to view the brain as a series of modules in dynamic equilibrium with each other, forming new connections, and responding to environmental stimulus [72]. If a module is damaged or suppressed, it may be able to regrow or restore that function, prompted by relatively simple stimuli.

The most common form of stroke is an ischemic stroke – caused by a blood clot that chokes off the affected area from normal blood flow. If treatment is not sought immediately to clear the blockage, the brain cells served by the affected arteries can be starved of oxygen and shut down, causing irreparable damage. In this case, the functionality provided by those dead cells has been lost, but it appears that it is possible to regrow that functionality in an adjacent part of the brain [21]. However, not all disability stems from damaged cells – there is also evidence to suggest that otherwise healthy cells that experience a period of disuse, caused by localized swelling, can adopt a form of learned paralysis, which can persist after the swelling has subsided. In this case, the brain only needs to reset the healthy cells to restore functionality, rather than re-growing it [72]. In both cases, mirror therapy and haptic stimulus may be able to help.

A.2.1 Simultaneous and Symmetrical Movement

The premise of mirror therapy is that by observing an unaffected limb in motion, the brain can map that stimulus to a damaged area controlling the affected limb, and use that stimulus to regrow function [3]. This is possible due to the presence and function of mirror neurons, which can fire when the user merely observes someone else performing a movement, and be involved in multiple types of function, including vision, motor commands, and proprioception. It is this ability to interact with both vision and motor commands that suggests why mirror therapy might be effective [72].

Multi-tasking is easier for humans when they receive simultaneous stimuli from different senses (such as vision and touch), rather than concurrent stimuli from a single channel, which will cause one to take priority [74,77]. This suggests that humans can process concurrent sensory data in large amounts, so long as it originates from different modalities [41,51]. It is suggested that mirror therapy might be made more effective if the patient received simultaneous proprioceptive feedback that matched the movement of the hand to the reflected image [12,62]. This referred sensation can be produced even in healthy subjects who observe tactile simulations, and seems to suggest that adding haptic stimulus to the existing mirror therapy protocol might enhance the mirror therapy effect [81].

Mirror therapy therefore engages in bimanual rehabilitation by using the patient's unaffected arm to aid their affected arm, which leads to the use of symmetric movements [54]. Malabet describes three types of symmetrical movement models: "joint space symmetry (JSS) where the motions are mirrored and the joints on each limb follow the same angles, visual symmetry (VS) where the hands move in the same Cartesian directions, and point mirror symmetry (PMS) where the hands rotate around an arbitrary point in space

[50].” This study used a method where a movement was performed by one hand across one degree of freedom, after which an attempt was made to mimic the movement with the affected hand; by which the researchers concluded that the visual symmetry space performed better.

McAmis and Reed performed a similar trial, which confirmed the results, though they found certain instances where the JSS space was the easier at higher rates of movement. They hypothesize that this may be due to the “neurological and biomechanical advantages of JSS appearing during high frequency tasks, such as when clapping one’s hands”, and conclude that a combination of VSS and JSS symmetry modes might be most effective for bimanual rehabilitation [54]. There is also evidence to suggest that movements that utilize both the hand and arm together will be more effective than movements that address the hand and arm separately [57].

A.2.2 Mirror Therapy Outcomes

Ramachandran and Altschuler are two of the originators of this treatment, introducing mirror visual feedback (MVF) in 1992, as a way to investigate a treatment for phantom pain in amputees [72]. Altshuler describes the function of the technique in his 1999 article, saying that “the mirror provides patients with “proper” visual input— the mirror reflection of the moving good arm looks like the affected arm moving correctly—and perhaps “substitutes” for the often decreased or absent proprioceptive input. At this early stage, they begin to amount evidence that mirror therapy can be beneficial for some stroke survivors and suggest further clinical trials. Ramachandran does mention that not all patients show the same degree of recovery, and suggests that this variability may be due to the exact location if the injury and duration of paralysis following the stroke [72].

These trials follow after a number of published case studies, and in 2007 Sütbeyaz et al conclude the first randomized controlled trial, testing mirror therapy efficacy for lower extremity hemiparesis following a stroke, and confirm that it “enhances lower-extremity motor recovery and motor functioning in subacute stroke patients [80].” Members of the same group follow up with a similar trial for upper-extremity hemiparesis the following year, and find similar results – noting an improvement in hand movement immediately after 4 weeks of treatment that sustains in comparison to a control group past six months [90]. A more recent controlled trial has similarly positive results, claiming that mirror therapy was effective, notably on the hand, wrist, shoulder, elbow, and forearm [44].

One study attempted to recreate the mirror therapy treatment inside a fMRI machine, in an attempt to isolate and identify the specific neural networks that are engaged by the treatment. They did confirm supplementary activation during the mirror therapy sessions in two visual areas of the brain – the right superior temporal gyrus, and the right superior occipital gyrus. However, they did not see similar activation in the frontoparietal mirror neuron system, which would question the role of mirror neurons in the mirror therapy effect [53]. However, they discuss that they were not able to obscure the affected hand during the experiment in the fMRI machine, instead asking the trial subjects to focus intently on the unaffected hand. Leaving the affected hand visible may have skewed the results, when compared with the rest of the trials that managed to hide it.

Another fMRI study focuses on robotic assisted systems, used in conjunction with virtual reality presentations of the mirrored data, and conclude that their system was able to optimize behavioral performance in the trials, and “selectively recruit targeted neural circuits [57].” Another virtual reality trial is less supportive, finding no apparent effect on motor performance through the use of the system [12]. Merians et al are careful to mention that robotic and virtual reality systems have not been investigated with the highest level of clinical trial evidence yet, as most published work has been limited to case studies, feasibility studies, or studies without control groups. It is noted that virtual reality systems are relatively rare and expensive, which may contribute to the paucity of high quality studies concerning them [3].

There is some discussion of mental imagery as an analogue for mirror therapy, which is essentially the patient picturing the moving limb in their imagination, rather than observing one in a mirror. To be effective, this requires intense training and rehearsal, but it has the potential to active the same neural circuits as the mirror (Ramachandran & Altschuler, 2009) [72]. However, Cacchio discounts mental imagery as effective, while he is quick to support mirror therapy [14].

It is widely accepted that mirror therapy, like other neurorehabilitation methods, will be most effective when it is employed within 30 days of the injury. Byl and Abrams conduct a study focused on stroke survivors beginning therapy after a period of six months following their stroke and conclude that even after that later starting point progress could be made. They found significant improvement in sensory discrimination, fine motor control and function independence, and note the need for "directed goal-oriented, repetitive, rewarded activities [13]."

A.2.3 Haptic Interventions

Haptic technology has been widely deployed in clinical settings using robots, force feedback devices and other large-scale implementations, but has yet to be fully tested in wearable configurations outside of the lab. One early study was conducted by Lieberman and Brazeal, in pursuit of a wearable robotic suit that could provide real-time corrective feedback to the user. Their goal was to use vibrotactile stimulus to teach students new motor skills more effectively than traditional methods. After constructing the arm of the suit and testing it in a 3D imaging studio, they were able to achieve a notable performance improvement of 15% over the control, with a 7% percent improvement in learning speed – figures which improved over certain joints that were more receptive to haptic communication [45].

One technical finding of note in the Lieberman and Brazeal study was their observation of the inefficacy of typical off-center vibrating motors, which take a long time to spin up, and only provide generalized rotation vibration feedback. The authors instead recommend using linear resonant actuators, which offer much improved response times and more directed linear stimulus.

A directly relevant study was conducted by Jiang et al in 2008, investigating the improvement of finger force control in patients with multiple sclerosis using haptic feedback. They constructed a mirrored finger setup, with the unaffected hand using force sensors to measure and send stimulus to vibration motors taped to the fingernails of the affected hand, with the goal of sending guiding haptic impulses to help control the standard task of lifting a glass of water [40]. They produced strong results, showing the use of the haptic system produced a 60% increase in efficacy in balancing the glass.

Of note is their production of two distinctive feedback modes – amplitude based feedback (ABF), which simply maps the force sensor input to vibrational intensity in each corresponding finger, and event-cue feedback (ECF), that instead offers corrective stimulus when one of the affected fingers slips out of balance from the rest. They were able to prove during their trial that subjects with more pronounced impairment preferred the ABF method, which provided more generalized haptic feedback, and subjects with less impairment preferred the more finely tuned ECF mode, which provides no haptic feedback when the subjects perform the tests correctly. This is a useful indicator of how a haptic system might be able to adapt to the changing needs of the user, as they make progress through their rehabilitation. Another technical finding they note is the low performance of a single force sensor per digit, and they show improved results with an averaged set of three force sensors on each finger.

A more recent study at Georgia Tech produces a haptic glove intended to demonstrate a Passive Haptic Rehabilitation effect, by teaching simple piano songs to patients with spinal cord injuries using haptic stimulus [51]. The study identifies the usefulness of music incorporated into physiological rehabilitation regimes, as the relative ability to perform each song is a reasonable indicator of overall performance and gives the patient a reliable indicator of their progress. Markow's group builds a wireless haptic glove capable of being used in conjunction with a piano keyboard and teaches a group of test participants a simple song, before testing their ability to play it back. Their findings show a statistically significant improvement in playback ability, when paired with a combination of audible and haptic feedback, and a statistically significant improvement in finger sensitivity over time, as measured by the Semmes-Weinstein test.

While Markow's study shows an effective implementation of a haptic glove in a rehab setting, it also provides valuable information about the design of haptic gloves for people with impaired hands. She specifically notes a preference for an open palm, as that helped the patients move their affected fingers into the glove, and the need for a rechargeable battery, as many patients would find changing batteries to be a challenging task.

A.2.4 Training at Home and Boredom

Malabet's paper observes that home training may "lead to faster progress and better results in motor relearning", but notes that patients are dissatisfied with their options for further training after they are discharged from rehabilitation [50]. McCabe notes that mirror therapy should become part of their planned exercise program with ''little and often'' being the mantra to follow [55]. One study mentions that they observed some of their participants exhibiting signs of boredom during mirror therapy, indicating there may opportunities for improvement with the protocol to improve efficacy [44].

A.3 Problem Statement

Home-based therapy offers significant benefits to stroke survivors who are struggling with the effects of hemiparesis. There is evidence that supplementing clinical treatment by performing exercises at home improves medical outcomes – including functional improvements and task-based improvements [22]. Home-based exercises can inhibit deterioration and prolong and enhance patient independence, which leads to improved quality of life [69].

However, despite these substantive benefits, patients do not always thrive under home-based therapy regimes, due to their different levels of adherence. Adherence can be affected by the patient's understanding of the exercise, their understanding of how their goals are set by the therapist, the nature of their symptoms, and their particular challenges with depression [5]. Given this context, many therapists report that getting patients to stick with a prescribed regime is notoriously difficult.

It is therefore appropriate to question what the elements of a successful home rehabilitation system would look like. Given the potential benefits, motivated stroke survivors should be able to maintain their own therapeutic program at home, but the lack of adherence first needs to be dealt with. One way of addressing this deficit is through the use of an interactive glove – designed to offer the therapist and patient a toolkit of interactions to enhance the at-home therapy regime.

An iterative design process was begun with the goal of creating a set of proof of concept models that proved the viability of the interaction concept and allowed for exploration of fabrication techniques. To accommodate a wide range of users, rapid prototyping techniques that could allow sizing adaptation were explored – such as the use of laser cutters and computer embroidery machines. A series of prototypes was then created, designed to show the potential for multi-modal interaction.

A.4 Glove prototype

Figure 114 – A glove prototype, showing pressure sensors, vibrotactors, and LEDs installed.

The initial proof of concept was designed to demonstrate the basic potential of the interaction model, allowing finger-mounted force sensors on one hand to trigger haptic vibration motors on the corresponding fingers of the other hand, shown in Figure 114. Earlier prototypes had provided an Arduino-based electronics platform to build on, so the focus of this stage was to create a new set of gloves that could be connected to it. To expedite the process, prefabricated gloves (the type used for coin and stamp sorting) were selected that offered a simple base to tack the requisite electronic components on to. Force Sensing Resistors (FSRs) from Interlink Electronics were selected to capture the pressure and tapping movement from each fingertip, and shaft-less coin vibration motors from Precision Microdrives were selected for the haptic feedback.

The components were each positioned on top of the gloves while being worn, placing the motors just below the fingernail of each finger on the left hand, and placing the contact point of the FSRs under the fingertip. Each component was fixed into place using double-sided tape, and then secured by using thread to tack it firmly onto the glove. The components were then connected back to the prototype breadboard using a series of jumper cables, and a second glove was slipped on top to protect the circuit layer.

Firmware was devised that took pressure input from the FSRs through the analog ports of the Arduino, trimmed the top and bottom of the range to eliminate unintentional input while being worn, and input from any accidental shorts. This input was then used to trigger the haptic motors whenever a value in range was detected. A serial connection was established to an adjacent laptop, which sent the same FSR input to a Processing sketch, which triggered a note from a piano scale, timed with the haptic motor. In this way, the right hand was able to trigger notes tapped on a desktop while wearing the glove, which played the notes simultaneously with audio and haptic feedback across the hand, mirrored from thumb to thumb, index finger to index finger, and so on.

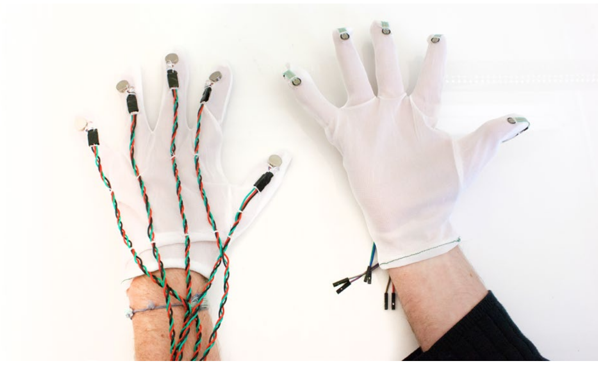


Figure 115 – Initial prototype, showing the wired circuits of the FSRs and coin motors.

This initial proof of concept was used to demonstrate the viability of the interaction, and the effectiveness of the hardware platform. It also demonstrated the necessity of improving the sensitivity of the components, their ease of positioning, and the need to improve the comfort and durability of the glove. A major challenge that emerged, however, was the length of time it took to calibrate the glove to each new user's hand. It was necessary to research new ways of construction the glove that would allow this process to run more quickly.

A.4.1 Finger Mounted Force Sensing Resistors

Force Sensing Resistors are well suited for use in this context, as they are thin and light, and adequately sensitive for their size. Yet the available commercial models are all inappropriate for use in the fingers of a glove, as they are designed with long plastic tails, spiked metal conductive prongs, and are not available in sizes and shaped that readily fit fingertips. For these reasons, it is difficult to sew them into the fingers of interactive gloves, and they are challenging to keep in position, as the sensing area is too small for most fingers.

It became necessary to investigate the design of a sensor more suited for use in gloves, one built using more comfortable fabric materials, based on a design that matched the strike zone of a finger tip under pressure, and one that allowed for the widest possible "sweet spot" that would reduce the number of erroneously ignored inputs. Building on earlier work by Leah Beuchley and members of her MIT research group, there are many examples of rudimentary FSRs built using conductive fabric and semi-conductive polymer, such as 3M's Velostat [67]. These designs work by placing two pieces of conductive fabric next to each other without touching on top of a piece of the semi-conductive polymer. In the default state, the pieces of fabric are not capable of transferring current from one another, as the resistance of the polymer is too high. As the polymer is placed under pressure, the resistance drops, and it is possible to complete the circuit, while measuring the relative change in resistance across the sensor. This is the basic interactive property that allows the sensor to function, but it can be improved to make more sensitive and finely tuned sensors.

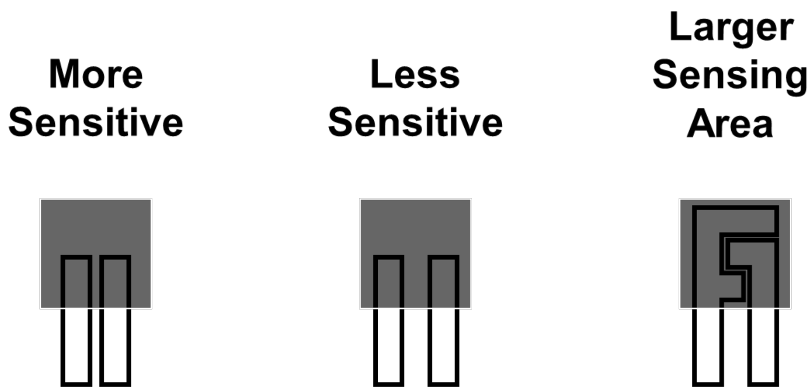


Figure 116 – Illustrations showing the effect of pattern on the sensitivity of DIY sensor design

The shorter the path taken through the semi-conductive polymer, the lower the resistance of the circuit – illustrated in Figure 116. This necessitates the design of a circuit pattern for the conductive fabric that has as small a gap as possible between the two sides. This also explains the default design of most commercial FSRs, which features an interdigitated pattern of small fingers on the conductive layer, which gives many opportunities for a short gap to be bridged, regardless of where the focal point of pressure is across the center. Replicating the small gap width of the commercial sensors is challenging with hand tools, so investigation was put into using a laser cutter to produce the conductive layers. Conductive fabric is resilient, but 3 passes of the cutter at a relatively high power and speed was effective at cutting through the pattern without scorching the fabric, and the gap width left by the laser was tight enough to allow for a highly responsive sensor.

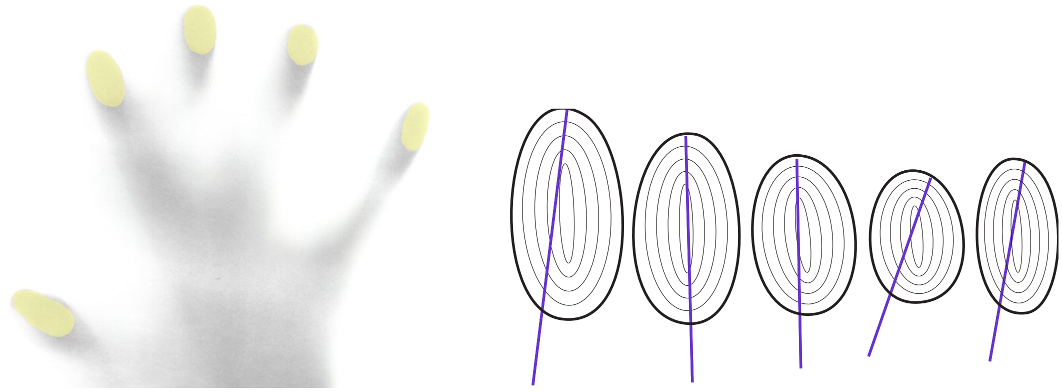


Figure 117 – Highlighted fingertips scan, and the orientation and layout of the initial fingertip FSR design.

Following this precedent, a design was created to match the outline of the sensors to best fit the fingertips of the author. This was accomplished by scanning fingertips on a flatbed scanner, with enough pressure to deform the fingertips into a flattened state. An outline was taken from this shape, illustrated in Figure 117, and translated into a sensor design that featured concentric interdigitated elements that followed the contours of the scanned shape.

Before cutting, conductive fabric was selected and ironed it onto a sheet of double-sided paper-backed adhesive. With one side bonded, the fabric was ready to be cut. Once the circuit shapes were finished, they were lifted off the cutting bed using the temporary adhesive of a Post-it Note to hold the spacing in place. Peeling the remaining paper backing off the circuit layer, the adhesive side was placed against the cut fabric back of the sensor, and the circuit layer was ironed and permanently attached to the backing. A matched shape was cut by hand out of the semi-conductive polymer and affixed to another fabric backing with double-sided fabric tape, as the polymer was sensitive to the heat required to bond

Wunder Under. With the two sides of the sensor completed, they were permanently bonded to each other using fabric glue. The complete set is shown in Figure 118.

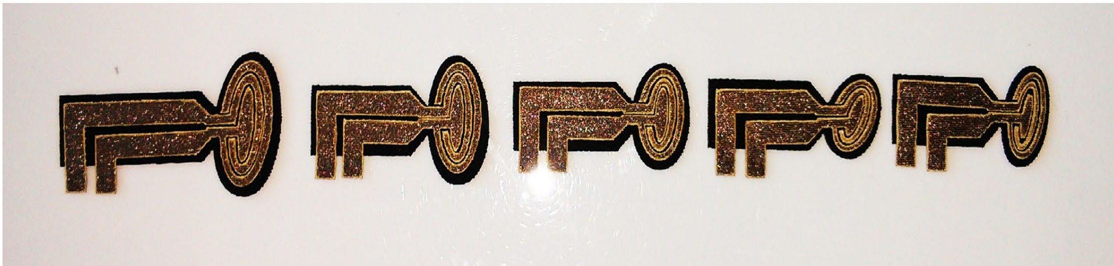


Figure 118 – The completed set of custom e-textile pressure sensors

These new sensors were tested and found to be highly responsive and stable in position. They were also comfortable and conformed to skin more readily that the plastic film varieties. This was first opportunity to observe how the design characteristics of specific components could alter the perceived encumbrance of the glove.

A.4.2 Linear Resonant Array motors

The high response rate of the sensors dictated that the motors should have a similar response profile, capable of translating a tap form one finger into a felt tap on the other. Conventional pancake motors are capable of generating strong vibration output, due to their eccentric oscillating masses, but they also take a relatively long time to start and stop, which makes them too indistinct and unresponsive for use in this case. LRA motors are powered by specialized AC drivers, which allow for precision control, which made them a suitable fit for this project. These drivers are only available in a tiny surface-mount package, which necessitated the development of a breadboard-scale breakout board to allow the LRA motors to be incorporated into the project.

The LRA motors were tested and adhered on to the fingers of the glove prototype, wrapping them in a fabric sleeve to help stabilize them and protect the connections. Once connected to the breadboard, the paired gloves were tested, and the LRA motors were highly responsive, communicating every distinct tap from the sensor glove to the corresponding fingers of the motor glove.

However, as the motors were able to provide useful haptic feedback, it was also clear that placing the motors directly on top of the fingertips would impair the movement and interaction of the wearer's fingers. This was a sub-optimal solution for a pair of gloves aimed at enhancing existing occupational therapy exercises, so the decision was made to mount the LRAs on the back of the fingers, below the nail. This cleared the motors out of the way, but later studies showed that positioning the motors on the dorsum of the hand led to interference in perceiving the individual pulses of each motor [78]. It was clear with this early exploration that there would be trade-offs in placing these vibrotactors, and the preference would need to be given for either physical object interaction, or haptic feedback.

A.4.3 LED sleeves

While the natural interaction of the glove is based around haptic feedback, LED indicators are useful for introducing the concept to the user and demonstrating the interaction to people who are not wearing the glove. For this model, each finger of the sensor glove was outfitted with a sewn Lilypad LED of a distinctive color, and the same color of LED was sewn onto the corresponding finger of the motor glove. In this way, shown in Figure 119, when a finger of the sensor glove was triggered, the matched color LEDs would light up on the fingers of both hands, demonstrating the connected nature of the concept.

Figure 119 – Final working proof of concept model – with motor glove on the left, sensor glove on the right.

A special sleeve was developed that was sized to slip on and hold the LEDs in place at the base of each finger. These sleeves were designed to provide a stable base to stitch the LED contact pads into the bare coiled ribbon cable strands, while allowing for some adjustment of the position of the LEDs for each user. To protect the LED components, they were sewn into a pocket on the sleeve, which allowed for the LEDs to shine through a laser cut pattern of tiny holes. This pattern was overlaid with an adhered layer of white fabric that acted a diffuser, spreading an even glow of light through the cuts.

A.5 Fit, donning and doffing

The fit of interactive gloves has a significant impact on the performance of the gloves with each user, due to the variability in the size and structure of each human hand. Fit helps to address both the comfort of the glove, as well as the accuracy of the sensors and effectors, which need precise positioning to be effective. This challenge was addressed by developing a system built around a parametric model that can structure custom patterns for the fabric, circuit layer, and component positioning, based off the hand measurements and requirements of each user.

Figure 120 – Three different closure mechanisms, to aid with donning and doffing of the glove.

Donning and doffing by users with paretic limbs is also a challenge, particularly those with hands in spasm, which led to the development of a series of prototypes – shown in Figure 120 – that explored various donning strategies. It is a key requirement that stroke survivor can easily manipulate the glove on their own, which requires the investigation of several novel closure mechanisms. A system was developed to look at ways to have the entire back or face of the hand open and reclose securely, allowing the user more direct access to the base of the fingers in the glove, which aids in the donning process.

A.6 Circuit density and printed circuits

Given the large number of components required, and the un-insulated nature of many e-textile circuit technologies, circuit density becomes a problem in the arrangement of components and the circuit traces that connect them. Many alternative circuit routing and printing technologies were considered and evaluated, that maintain sufficient electrical performance, while allowing for unrestricted movement of the hand.

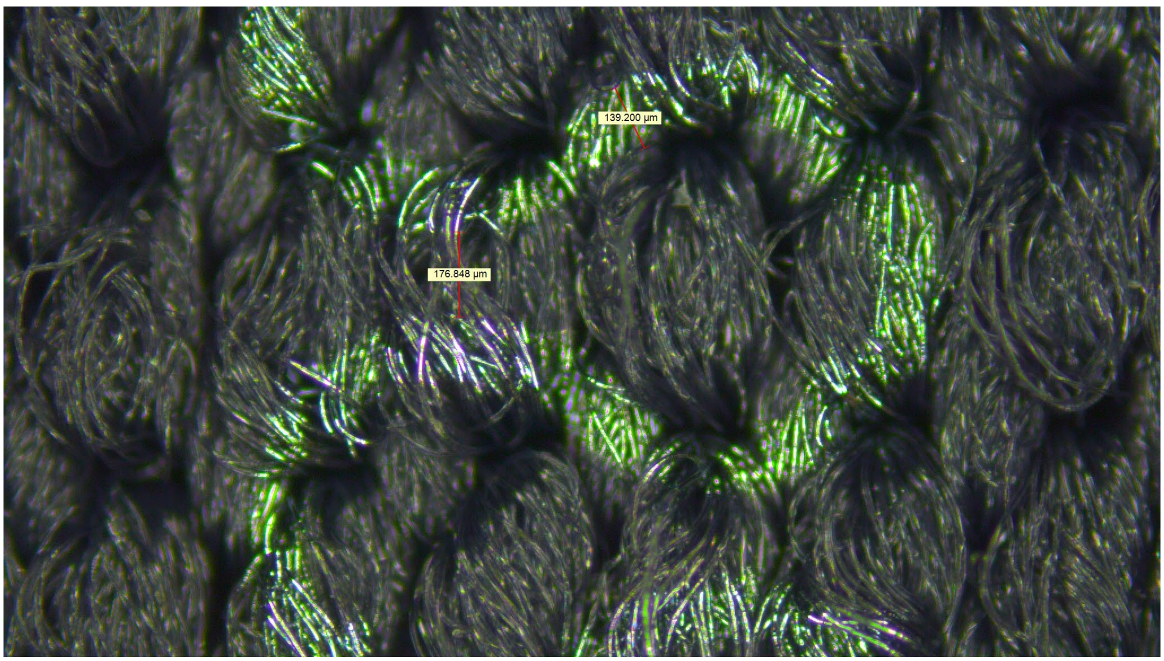


Figure 121 – A serpentine circuit pattern deposited on a knit nylon fabric, showing a 176.8 micron gap.

To do this, preliminary research was started on using an emerging technology that allows for the vapor depositing of circuits directly onto the surface of the glove. This pattern is shown in place on the fabric fibers in Figure 121. Using the parametric model described earlier with this new technology may allow for the production of individualized circuit traces that conform to the hand of each user and preserve registration with the fabric cut pattern.

A.7 Stretch Circuit Characterization

One of the more difficult problems still facing e-textile designers is the integration of circuits into knit fabrics. Knits are popular with garment producers because they move and stretch easily – due to expandable linked loops of the knit stitch, and the frequent inclusion of Spandex or Elastane fibers. This makes knit fabrics an ideal match for the many activity tracking garments that are beginning to make their way to the market. However, the integration of functional circuits into these activewear garments can be non-trivial problem to overcome –especially in a way that preserves the unique stretchable properties of the knit fabric.

While there are many options for integrating e-textile circuits into fabric, a sewing machine equipped with conductive thread offers one of the most versatile, available, and repeatable solutions. Stainless steel thread has recently risen in popularity – due to its conductivity, durability, and resistance against oxidation [32]. Yet stainless steel thread can be challenging to integrate onto a knit fabric substrate, as it is typically inelastic in its construction. This lack of elasticity can create stress along the stitches while the garment is in use, and potentially lead to the degradation or breaking of the conductive fiber.

This study examined the performance of two popular varieties of stainless-steel conductive thread in use as machine-sewn conductive traces in stretchable e-textile garment applications. To test these threads, a set of standardized fabric swatches were prepared, each stitched with one of three common stitch patterns for each of the threads in the experiment and measured the changing strain and electrical resistance under controlled load. The resulting data was then examined to identify any interaction between thread composition and stitch type in the durability, fabric stretch behavior and conductivity of the thread under variable load. This led to a recommendation of the best performing combination of thread and stitch pattern to make viable e-circuit traces on stretch substrates.

A.7.1 Background

The threads used in this experiment are commonly available from prominent electronics retailers. Both threads selected are spun from filaments of Grade 316L stainless steel – a low-carbon, molybdenum-bearing alloy. The most notable difference between them is the structure of the yarn they are formed out of – one is based around long smooth filament strands, while the other is based on shorter staple fibers that have been spun into the yarn. Suzanne Watkins and Lucy Dunne note in “Functional Clothing Design” (2015), that “the ways in which fibers are formed into yarns may be as critical or even more critical to the final characteristics of the fabric than the type of fiber used” [87]. This observation helped guide the intention for this experiment – to examine the possible role of thread composition in e-textile performance outcomes.

Many researchers have looked at the performance of conductive thread in e-textile applications. Post et al examined multiple applications for conductive thread on denim fabric and noted that continuous filament fibers translate stress along their entire length until failure, while staple fibers can dissipate some force through a limited stretch. They also note a mean weld strength of the soldered connections used in their work (36.8N), which gives us a usable force measurement to target – 50N was selected as the maximum, to examine the range of forces a typical garment might experience [68]. Margaret Orth identified the "general characteristics of machine sewable yarns" and conducted a subsequent stretch experiment. Following this test, Orth noted that a 100% stainless steel fiber, after being loaded to 9N was unsuitable for sewing [64]. This prompted us to test currently available stainless-steel threads to see if they would perform in a similar way.

Researchers have also examined stretch circuits –particularly in the application of creating sensors. In 2007, Huang et al. measured the electrical resistance of conductive yarns – evaluating for integration into yarn-based piezo-resistive sensors. During the experiment, they characterized the linearity of the resistance/stretch curves – finding that "double wrapping" the thread provides a more linear curve [36]. In 2012, Gioberto and Dunne characterized the resistance and elongation of a top thread cover stitch conductive thread trace. They noted the utility of a looped conductor pattern in allowing the stitch to stretch with the elastic fabric, and the essential role that the fabric substrate played in the characteristics of the sensor [25]. This works therefore leads to a continued search for any relationship between the fabric substrate, top stitch, and conductive yarn stitch in the performance of the conductor.

A.7.2 Methods

The intent of this experiment was to evaluate the mechanical and electrical performance of two types of conductive threads sewn with various types of stitches into swatches of knit fabric. The independent variables in this experiment included the type of thread used, the type of stitch used, and the amount of force used to linearly stretch the knit swatch. Two types of thread were evaluated: a staple fiber stainless steel, and a filament stainless steel. Three common stitches were tested: straight stitch, regular zigzag stitch, and step zigzag stitch. A step zigzag is different from a regular zigzag stitch as it has additional top thread stitches anchoring the bottom thread. The dependent measures consisted of the total length of the experimental swatches and electrical resistance of the thread under load – which were measured for each combination of independent conditions.

A.7.3 Construction of Swatches

To construct the swatches, a common consumer-grade sewing machine was selected – a Brother PC420PRW model, using the standard presser foot provided with the machine. The machine was installed with the provided 75/11 sharp needle and threaded with a top spool of black polyester thread and a lower bobbin of conductive thread. Every swatch was cut from the same piece of fabric – a standard black one-way stretch cotton Lycra knit, chosen due to its prolific use in athletic wear production. The fabric was cut into 101.6 mm x 350 mm rectangles with the stretch in the long direction. The conductive traces in each sample were 200mm long, with 5 mm of back stitching to anchor the thread on either end. A folded-over pocket was fashioned with a 15 mm length on each end, spaced 20 mm away from the end of each trace.

Figure 122 – From L-R: Staple fiber straight stitch, staple fiber zigzag stitch, filament thread step zigzag stitch.

Two conductive threads were selected for this test, as shown in Figure 122. The first was procured from Adafruit.com – a two-ply stainless-steel thread listed at 52.49 Ohms/m. This thread is smooth continuous filament thread and will be referred to as “filament thread”. Our second thread was procured from Sparkfun.com – a spun stainless steel yarn listed at 91.86 Ohms/m. This thread is a textured staple fiber and will be referred to as “staple fiber thread”. Both threads are composed of 100% Grade 316L stainless steel with no non-conductive materials. After a brief period of experimentation, stitch length/width settings were selected as follows: Straight stitch at -/2.5, Zigzag at 3.5/2.0, and Step Zigzag at 4/1. Each was sewn at a moderate tension (selected as 4 on the machine).

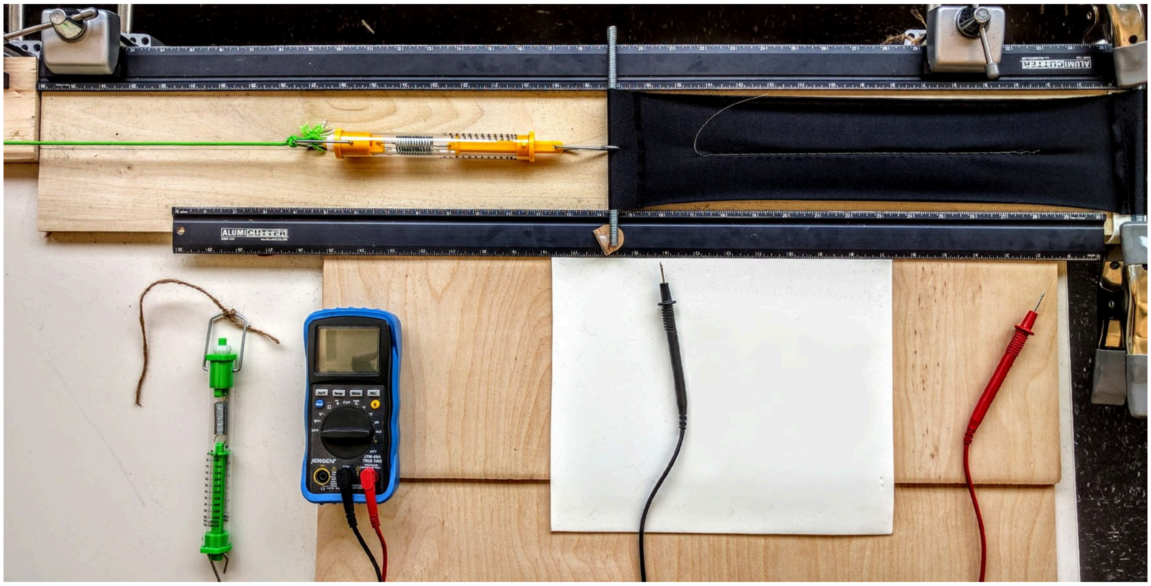


Figure 123 – The testing apparatus, showing the stretching mechanism, strain gauge, and multi-meter.

A.7.4 Apparatus

A desktop apparatus, shown in Figure 123, was designed to allow each fabric sample to be stretched and measured. During testing, a metal dowel was slid into each pocket at the ends of the fabric swatches. One dowel was clamped onto the end of the apparatus such that the edge of the fabric was secured at the zero mark on a measuring stick. An SI Manufacturing mechanical force gauge was hooked around the dowel at the opposing end of the swatch through the slit in the middle of the pocket. The gauge was pulled with a string to apply linear force to the sample. When a desired force was reached the string was secured with a clamp to lock in the applied force. A Jensen JTM-69A multimeter was used to take resistance measurements across the thread.

A.7.5 Procedure

Each sample was mounted to the apparatus as described above. The sample was then tensioned to a baseline force of 0.5N, then to the designated force for the experiment, and then returned to the baseline. The baseline force of 0.5N was chosen because it provided a stable and repeatable way of measuring the fabric at rest. Each experiment started at 10N of force and increased by steps of 10N up to a final test of 50N of force, or the point at which the thread broke. At each force reading the length of the fabric swatch was recorded. Three end-to-end electrical resistance measurements were also taken at each force checkpoint. If a thread broke during the tensioning process, the force at which the break occurred was recorded. The broken sample remained tensioned until a length measurement was taken and an electrical failure in the thread was identified. One final length reading was taken for the baseline.

A.7.6 Results

The straight stitches all broke in both threads – 5 of the staple fiber threads breaking between 8-11N, and 5 of the filament threads breaking between 19-34N. Puckering across the fabric was seen in the straight stitches at 10N.

Table 11 – Thread Breaks & Break Force. The Staple Fiber Step / ZigZag combination and the Filament / Reg. ZigZag both survived forces up to 50N.

Sample	# of Breaks	Mean Break Force/N
Staple Fiber Straight	5	9.2
Staple Fiber Reg. ZigZag	5	28
Staple Fiber Step ZigZag	0	stopped at 50N
Filament Straight	5	26.2
Filament Reg. ZigZag	0	stopped at 50N
Filament Step ZigZag	5	43.25

The staple fiber threads in the Regular ZigZag stitches started to straighten out between 0.5-10N and all snapped between 24-39N. Sections of thread in the Step ZigZag pulled taught after 30N but reached 50N with no breaks.

Table 12 – Mean Electrical Resistance of tested thread. The filament fibers demonstrated lower resistance in combination with all stitch patterns.

Fiber / Stitch pattern	Mean Resistance /Ohms
Staple Fiber Straight	22.01
Staple Fiber Reg. ZigZag	28.42
Staple Fiber Step ZigZag	24.07
Filament Straight	8.71
Filament Reg. ZigZag	16.97
Filament Step ZigZag	11.68

With the filament thread, in the Regular ZigZag samples, pronounced fabric puckering was observed after 20N, yet there were no breaks in any sample up to 50N. For the Step ZigZag, all samples broke between 40-50N. Once stitches broke, the thread seemed to release tension and settle at a lower value of force – with an average drop of 14.75N.

Factorial ANOVA was carried out to compare the resistance for the three stitch types under different forces for both thread types. No significant difference in electrical resistance was observed for different forces. Significant difference in resistance was observed between ZigZag and Straight stitches for both the thread types (p value<0.005).

Factorial ANOVA was carried out to compare the effects of Thread Type, Stitch Type and Force on the Stretch of the fabric:

Table 13 – The interaction effect of Force*Stitch (p value=0.012, <0.05) and Stitch*Thread (p-value= 0.000, <0.05), had a significant effect on the fabric stretch.

Tests of Between-Subjects Effects

Dependent Variable: Stretch

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	320.664[a]	29	11.057	49.553	.000	.935
Intercept	75344.269	1	75344.269	337654.595	.000	1.000
Force	176.819	5	35.364	158.483	.000	.888
Stitch	12.693	2	6.346	28.441	.000	.363
Thread	1.495	1	1.495	6.702	.011	.063
Force * Stitch	4.680	8	.585	2.622	.012	.173
Force * Thread	.779	5	.156	.698	.626	.034
Stitch * Thread	16.805	2	8.402	37.656	.000	.430
Force * Stitch * Thread	.472	6	.079	.353	.907	.021
Error	22.314	100	.223			
Total	135560.778	130				
Corrected Total	342.978	129				

a. R Squared = .935 (Adjusted R Squared = .916)

A.7.7 Discussion

The straight stitches all broke under load. This appears to be due to the difference in stretch ratios between the fabric substrate and the conductive thread. With the straight stitch having already elongated the conductive thread with no strain relief, the force has nowhere to go but into the fiber. The staple fiber thread seems to stretch and gradually pull apart under load, where the filament thread distributes the force along the length of the fiber until it snaps. These findings illustrate the utility of the zigzag stitch, as it provides for additional strain relief, delays the final elongation of the thread, and seems to more effectively spread the applied force into the fabric substrate through the top stitches.

The results obtained in the resistance analysis can be explained by the higher amount of thread required for the zigzag stitches compared to the straight stitches. The difference in resistance observed between the Regular ZigZag and Step ZigZag was not significant. There was no significant deviation observed from the manufacturer's reported resistance for either thread. It seems that without the thread making contact with itself across the stitch, the total resistance from edge to edge will be a simple product of total length of the thread in the stitch – it is not influenced by the pattern of the stitch.

Our findings indicate a significant interaction effect of Force & Stitch pattern, and Stitch pattern & Thread composition in the stretch behavior of the fabric. Upon subjecting the samples with varying levels of tensile force, it was observed that the Filament Fiber Regular ZigZag samples stretched to increased lengths and showed greater variability in the resting swatch length when compared to Staple Fiber Step Zigzag samples. This indicates that when selecting the thread and stitch type, consideration for the desired effect on fabric behavior needs to be made. Also, it would be advisable to ascertain the extent of force a fabric is expected to endure prior to selecting the stitch type to use. This could be especially useful in maintaining the desired level of stitch conductance and elastic properties of the fabric in use.

Finally, the two thread/stitch combinations that successfully sustained 50N were compared. The most obvious difference between the two zigzag patterns is the addition of the extra top thread stitches to Step Zigzag stitch. This seems to indicate a relationship between the number of top stitches in the pattern and the fiber composition of the thread in the outcomes that were measured.

Figure 124 – Detail of filament thread. Left: Intact thread in zigzag stitch. Right: Broken thread in step zigzag stitch.

The staple fiber/step zigzag combination may have been successful because the step zigzag seems to provide more points of control from the top thread over the bottom thread. For the staple fiber thread, this seemed to hold the fibers together under elongation – with the staple fibers exhibiting more stretchy characteristics as they stay entangled with each other.

By contrast the filament thread survived the 50N force with the Regular Zigzag, shown in Figure 124, but broke with the Step Zigzag. Under the microscope it can be observed that the Step Zigzag better controlled the filament thread – pulling it into a conventional angular Zigzag pattern. The Regular Zigzag combination had less control over the thread, allowing it to flow into open loops in a serpentine pattern. This seems to have allowed for the filament to survive the 50N force – there is more thread in the open loop pattern, providing more strain relief. The Step Zigzag, by better controlling the filament thread, incorporates less of the thread into the stitch. This seems to be supported by the mean resistance values of the two stitches at rest (Regular Zigzag – 16.97 Ohms, Step Zigzag – 11.68 Ohms), which were established had a linear relationship between resistance and length for both threads.

A.7.8 Conclusion

These findings suggest that there is a role for both filament and staple fiber stainless steel thread in the creation of machine sewn e-textile conductive traces – so long as they are combined with an appropriate stitch and consideration is given to how strain will be distributed along the thread. From these observations, filament thread seems to have better conductivity, and can sustain higher forces, while being harder to control with the machine. Staple Fiber seems to have lessened conductivity, and breaks more easily under smaller loads, while being easier to control.

While more work is needed, it can be confirmed that there a significant effect between Force * Stitch, and Stitch * Thread composition in the overall behaviour of the stretch fabric. The following recommendation can also be made – if you want to durably integrate a filament thread into a stretch fabric, use a regular zigzag stitch – if you have staple fiber thread, use a step zigzag. These combinations are the most durable during flexion, which enables a greater freedom of hand movement than those with greater physical resistance. This study was completed by the author, with the assistance of Reema Upadhyaya, Harrison Daniels, and Caity Taylor.

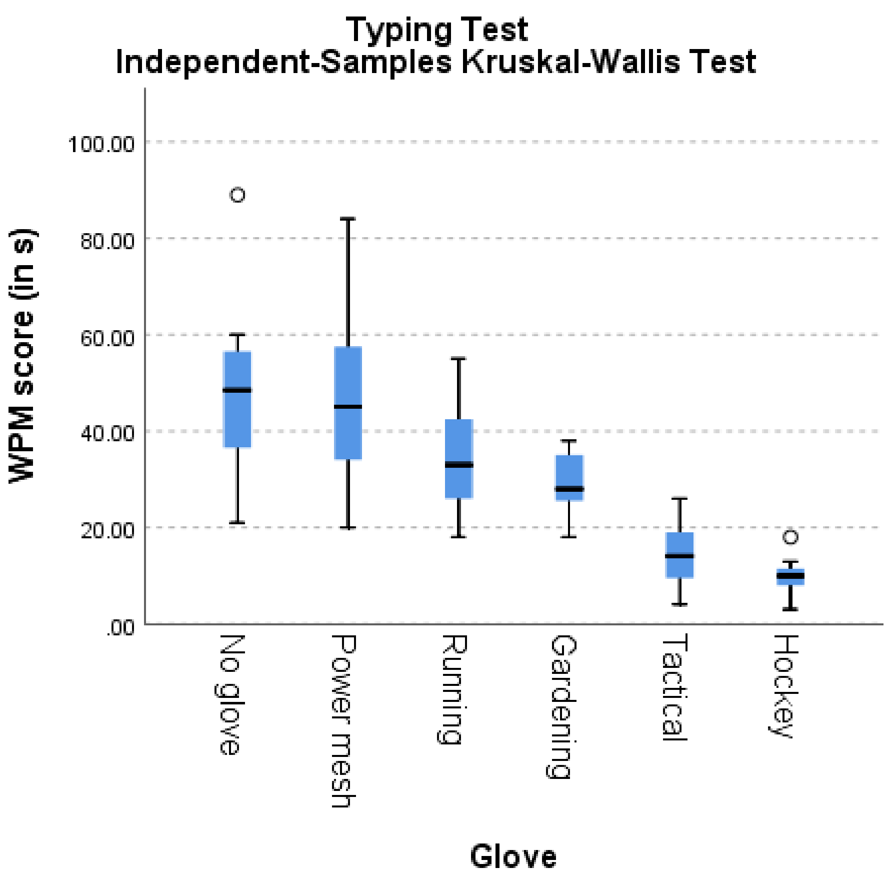
APPENDIX B. STUDY 1: GLOVE ENCUMBRANCE STUDY DATA

B.1 Task performance data

B.1.1 Typing test

Independent-Samples Kruskal-Wallis Test Summary	
Total N	72
Test Statistic	51.104[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.

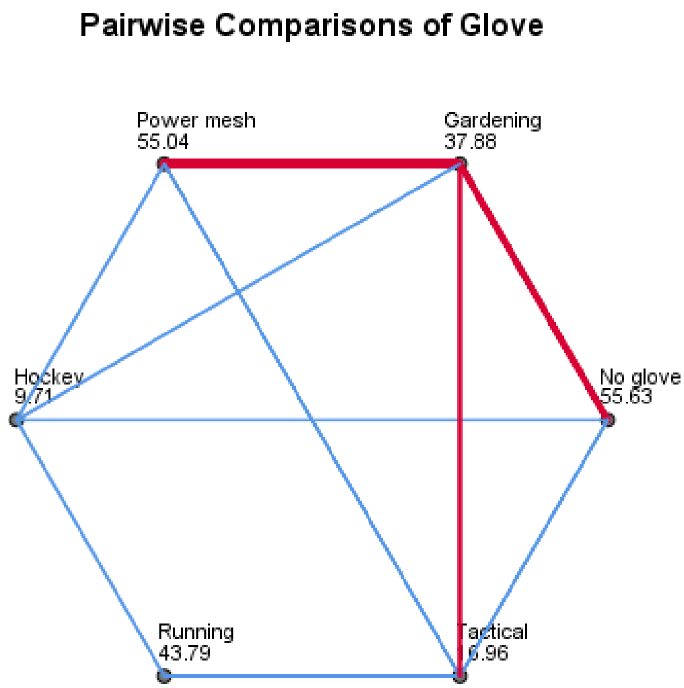


Pairwise Comparisons of Glove					
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Hockey-Tactical	7.250	8.542	.849	.396	1.000
Hockey-Gardening	28.167	8.542	3.298	.001	.015
Hockey-Running	34.083	8.542	3.990	.000	.001
Hockey-Power mesh	45.333	8.542	5.307	.000	.000
Hockey-No glove	45.917	8.542	5.376	.000	.000
Tactical-Gardening	20.917	8.542	2.449	.014	.215
Tactical-Running	26.833	8.542	3.141	.002	.025
Tactical-Power mesh	38.083	8.542	4.459	.000	.000
Tactical-No glove	38.667	8.542	4.527	.000	.000
Gardening-Running	5.917	8.542	.693	.489	1.000
Gardening-Power mesh	17.167	8.542	2.010	.044	.667
Gardening-No glove	17.750	8.542	2.078	.038	.566
Running-Power mesh	11.250	8.542	1.317	.188	1.000
Running-No glove	11.833	8.542	1.385	.166	1.000
Power mesh-No glove	.583	8.542	.068	.946	1.000

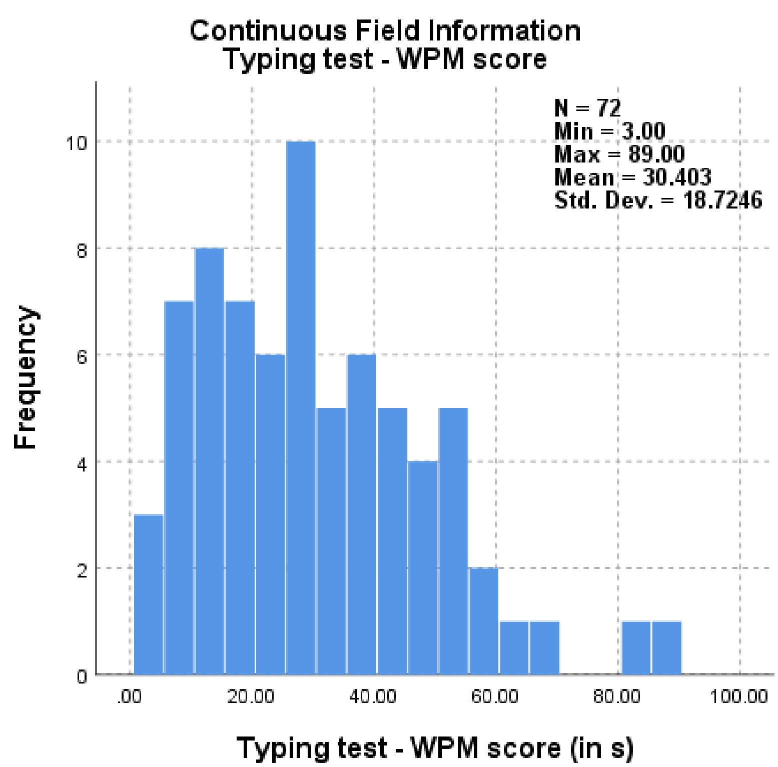
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



Each node shows the sample average rank of Glove.

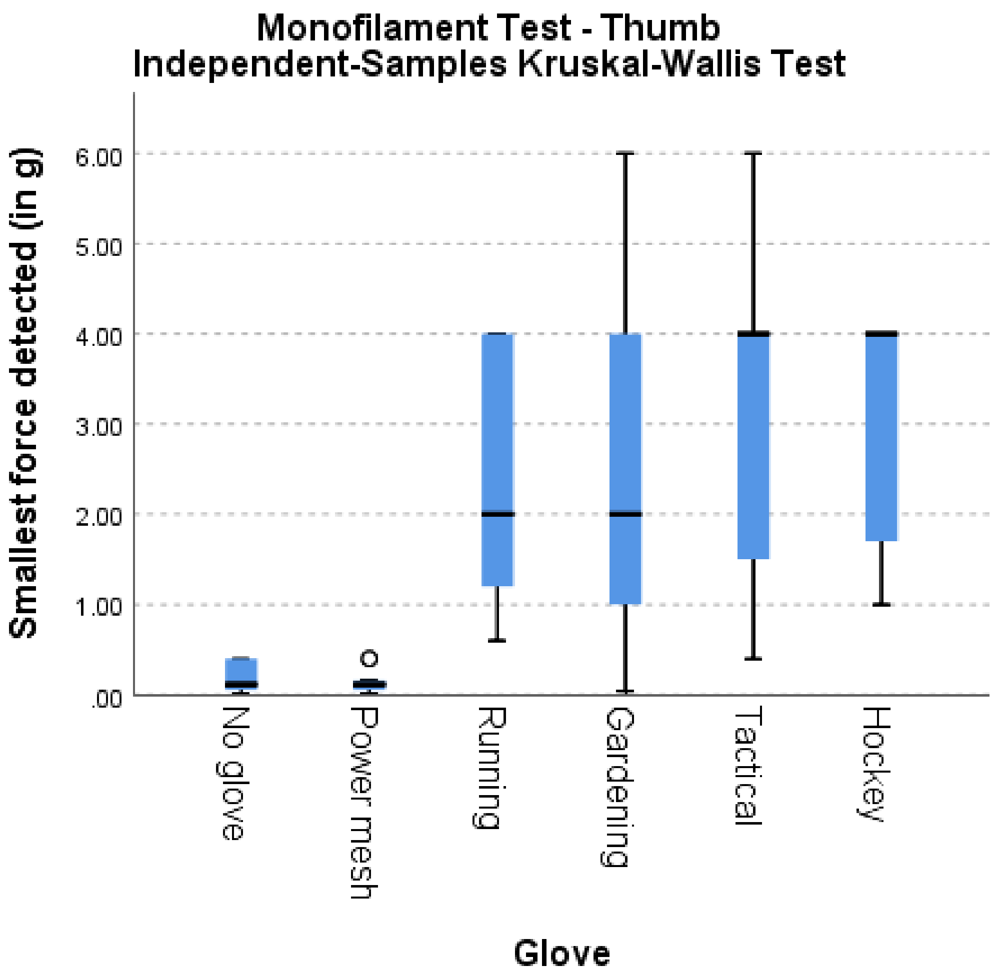


B.1.2 Monofilament test - Thumb

Independent-Samples Kruskal-Wallis Test Summary

Total N	72
Test Statistic	45.961[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

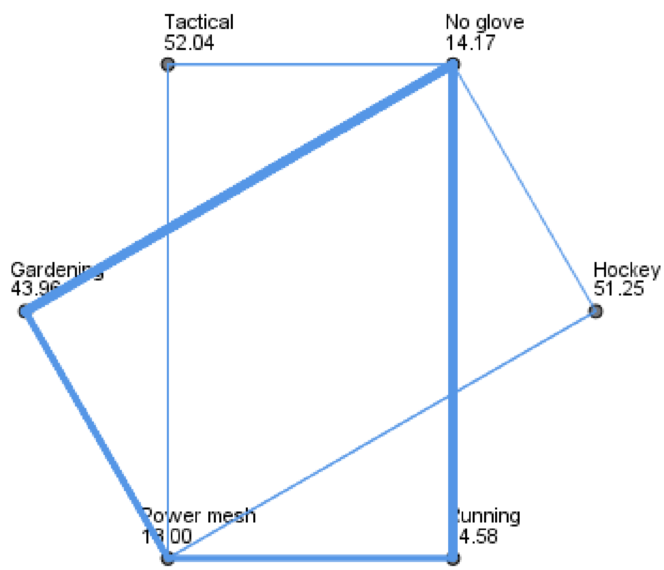
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Power mesh-No glove	1.167	8.425	.138	.890	1.000
Power mesh-Gardening	-30.958	8.425	-3.675	.000	.004
Power mesh-Running	-31.583	8.425	-3.749	.000	.003
Power mesh-Hockey	-38.250	8.425	-4.540	.000	.000
Power mesh-Tactical	-39.042	8.425	-4.634	.000	.000
No glove-Gardening	-29.792	8.425	-3.536	.000	.006
No glove-Running	-30.417	8.425	-3.610	.000	.005
No glove-Hockey	-37.083	8.425	-4.402	.000	.000
No glove-Tactical	-37.875	8.425	-4.496	.000	.000
Gardening-Running	.625	8.425	.074	.941	1.000
Gardening-Hockey	-7.292	8.425	-.865	.387	1.000
Gardening-Tactical	-8.083	8.425	-.959	.337	1.000
Running-Hockey	-6.667	8.425	-.791	.429	1.000
Running-Tactical	-7.458	8.425	-.885	.376	1.000
Hockey-Tactical	.792	8.425	.094	.925	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

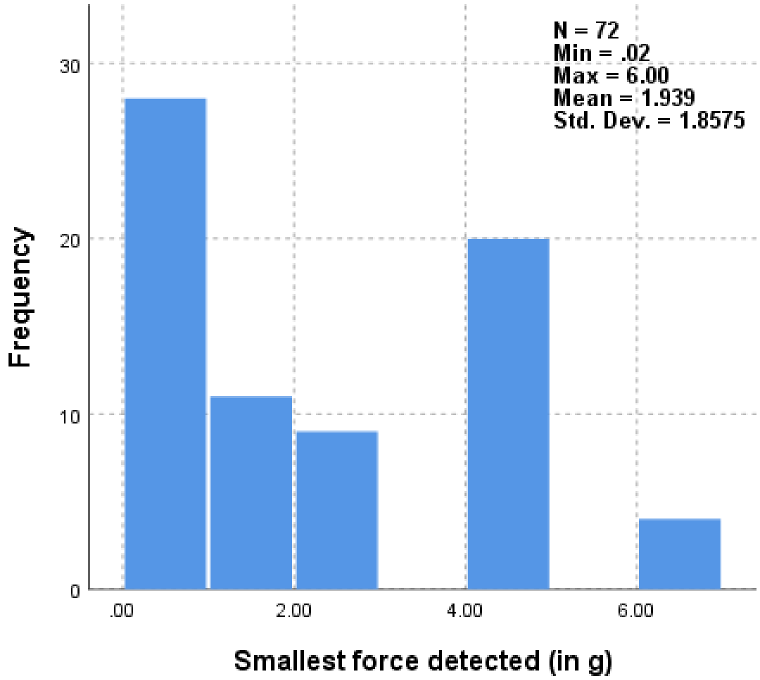
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.

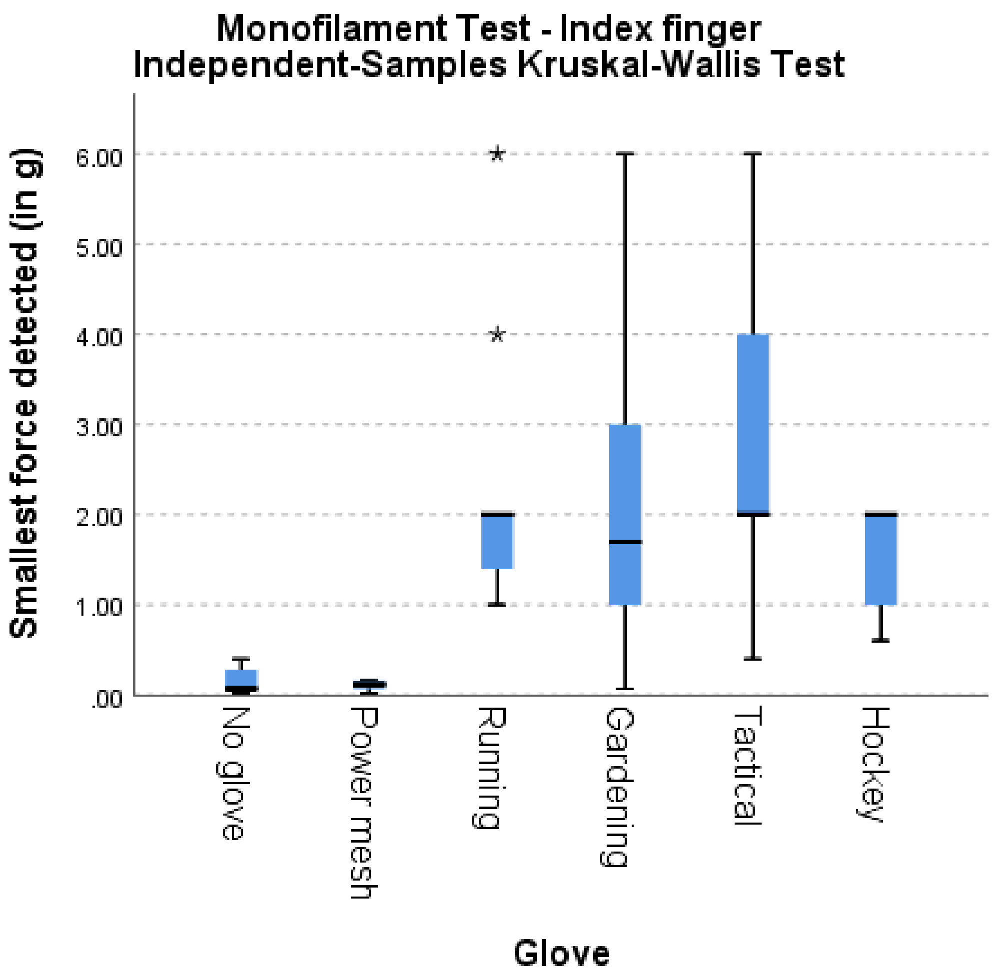
Continuous Field Information
Monofilament Test - Thumb



B.1.3 Monofilament test – Index finger

Independent-Samples Kruskal-Wallis Test Summary	
Total N	72
Test Statistic	47.282[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



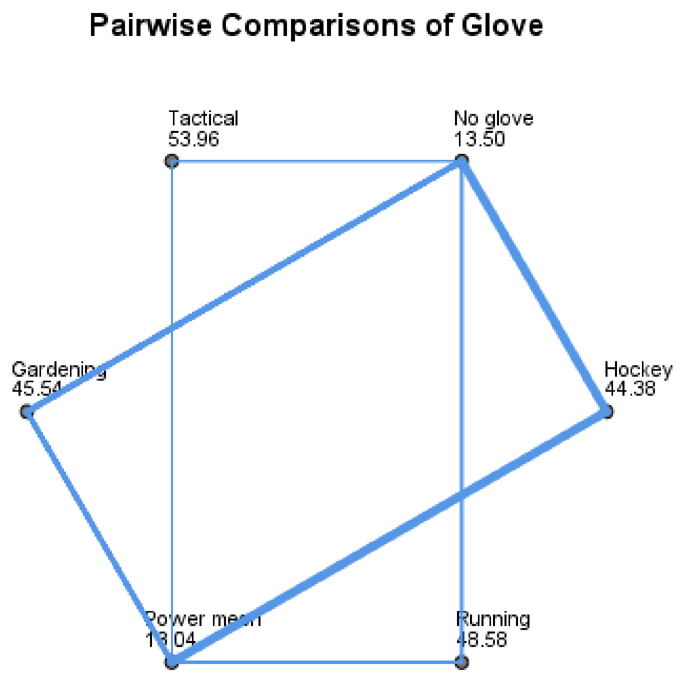
Pairwise Comparisons of Glove

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Power mesh-No glove	.458	8.414	.054	.957	1.000
Power mesh-Hockey	-31.333	8.414	-3.724	.000	.003
Power mesh-Gardening	-32.500	8.414	-3.862	.000	.002
Power mesh-Running	-35.542	8.414	-4.224	.000	.000
Power mesh-Tactical	-40.917	8.414	-4.863	.000	.000
No glove-Hockey	-30.875	8.414	-3.669	.000	.004
No glove-Gardening	-32.042	8.414	-3.808	.000	.002
No glove-Running	-35.083	8.414	-4.169	.000	.000
No glove-Tactical	-40.458	8.414	-4.808	.000	.000
Hockey-Gardening	1.167	8.414	.139	.890	1.000
Hockey-Running	4.208	8.414	.500	.617	1.000
Hockey-Tactical	9.583	8.414	1.139	.255	1.000
Gardening-Running	3.042	8.414	.361	.718	1.000
Gardening-Tactical	-8.417	8.414	-1.000	.317	1.000
Running-Tactical	-5.375	8.414	-.639	.523	1.000

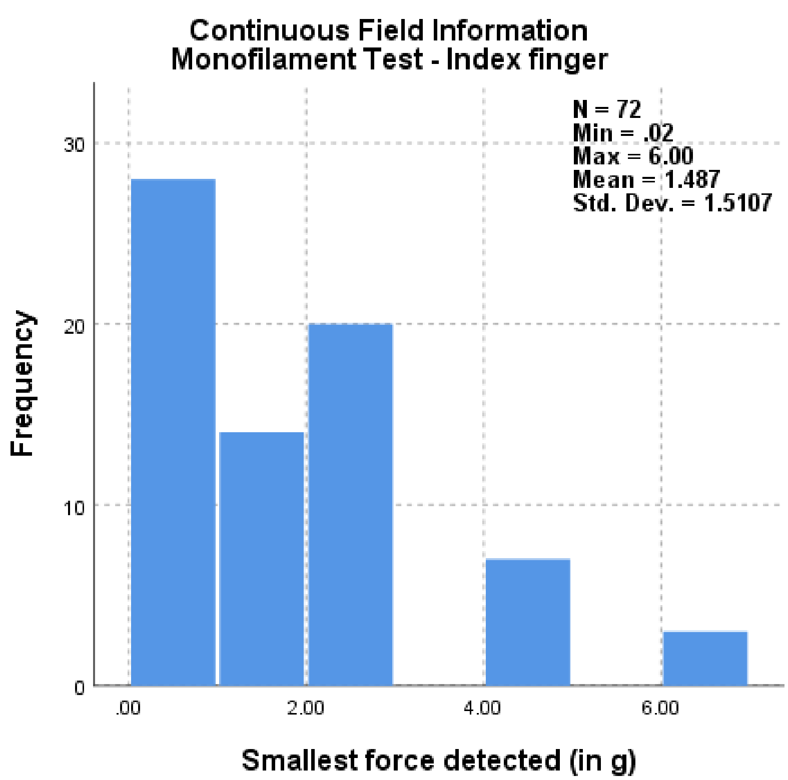
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



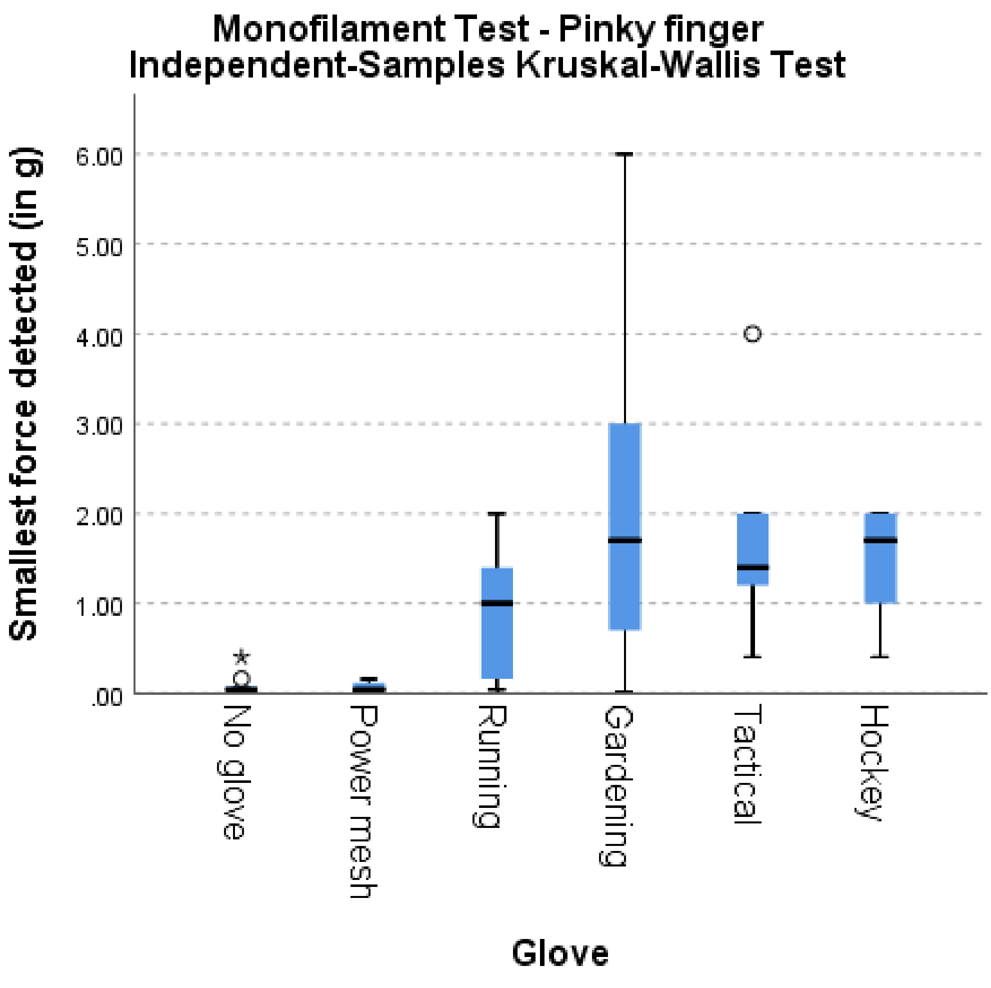
Each node shows the sample average rank of Glove.



B.1.4 Monofilament test – Pinky finger

Independent-Samples Kruskal-Wallis Test Summary	
Total N	72
Test Statistic	43.359[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



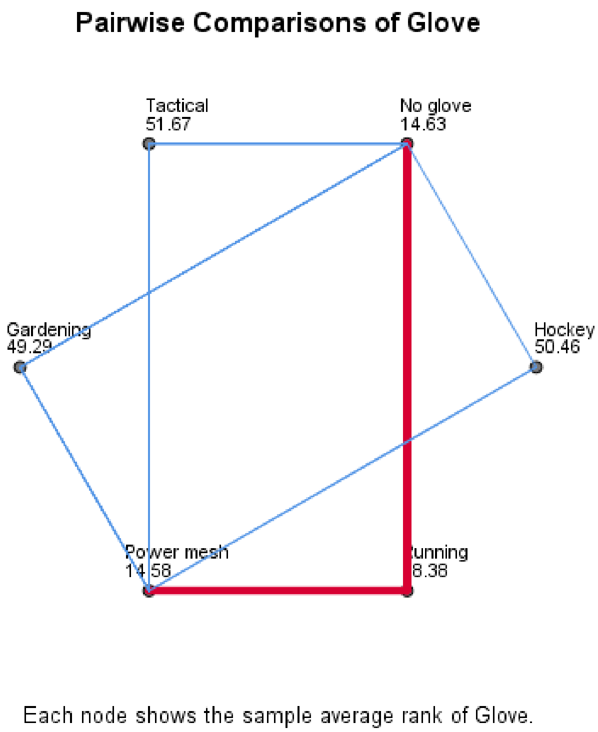
Pairwise Comparisons of Glove

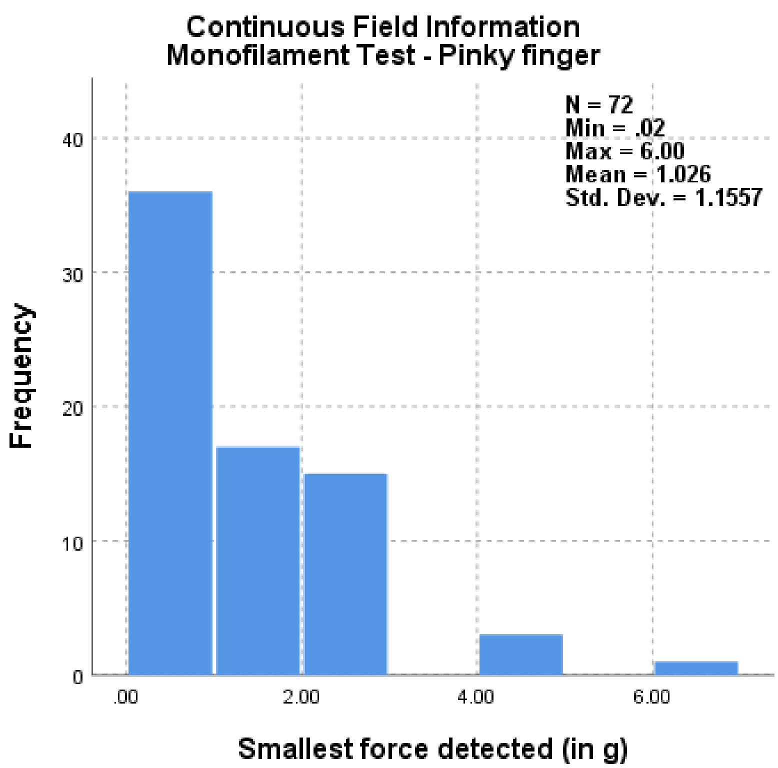
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Power mesh-No glove	.042	8.458	.005	.996	1.000
Power mesh-Running	-23.792	8.458	-2.813	.005	.074
Power mesh-Gardening	-34.708	8.458	-4.104	.000	.001
Power mesh-Hockey	-35.875	8.458	-4.242	.000	.000
Power mesh-Tactical	-37.083	8.458	-4.384	.000	.000
No glove-Running	-23.750	8.458	-2.808	.005	.075
No glove-Gardening	-34.667	8.458	-4.099	.000	.001
No glove-Hockey	-35.833	8.458	-4.237	.000	.000
No glove-Tactical	-37.042	8.458	-4.380	.000	.000
Running-Gardening	-10.917	8.458	-1.291	.197	1.000
Running-Hockey	-12.083	8.458	-1.429	.153	1.000
Running-Tactical	-13.292	8.458	-1.572	.116	1.000
Gardening-Hockey	-1.167	8.458	-.138	.890	1.000
Gardening-Tactical	-2.375	8.458	-.281	.779	1.000
Hockey-Tactical	1.208	8.458	.143	.886	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

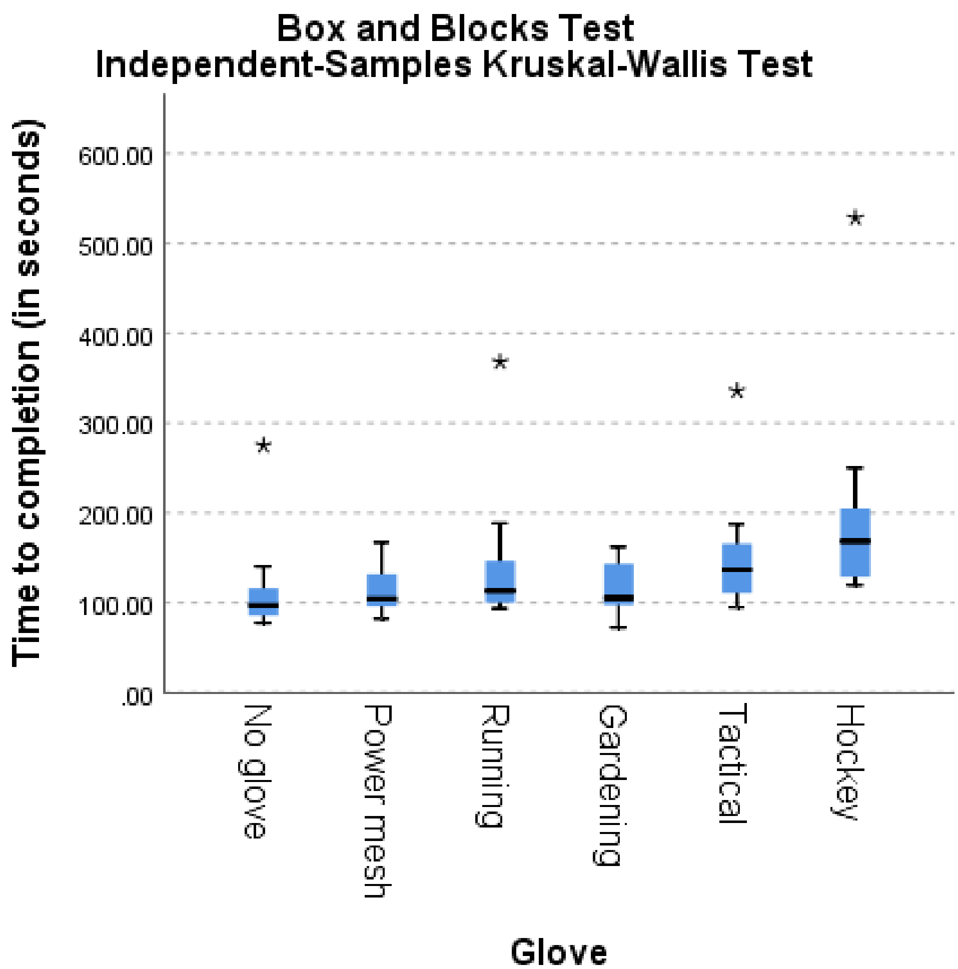

Pairwise Comparisons of Glove
Tactical
51.67
No glove
14.63
Gardening
49.29
Hockey
50.46
Power mesh
14.58
Running
8.38
Each node shows the sample average rank of Glove.


Continuous Field Information
Monofilament Test - Pinky finger
N = 72
Min = .02
Max = 6.00
Mean = 1.026
Std. Dev. = 1.1557
Frequency
0
10
20
30
40
.00
2.00
4.00
6.00
Smallest force detected (in g)

B.1.5 Box and Blocks test

Independent-Samples Kruskal-Wallis Test Summary	
Total N	72
Test Statistic	21.158[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.001

a. The test statistic is adjusted for ties.



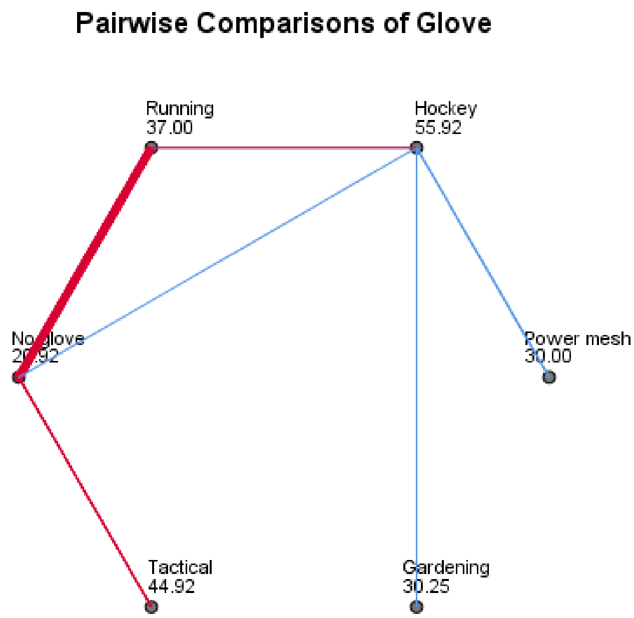
Pairwise Comparisons of Glove

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
No glove-Power mesh	-9.083	8.544	-1.063	.288	1.000
No glove-Gardening	-9.333	8.544	-1.092	.275	1.000
No glove-Running	-16.083	8.544	-1.882	.060	.897
No glove-Tactical	-24.000	8.544	-2.809	.005	.075
No glove-Hockey	-35.000	8.544	-4.096	.000	.001
Power mesh-Gardening	-.250	8.544	-.029	.977	1.000
Power mesh-Running	-7.000	8.544	-.819	.413	1.000
Power mesh-Tactical	-14.917	8.544	-1.746	.081	1.000
Power mesh-Hockey	-25.917	8.544	-3.033	.002	.036
Gardening-Running	6.750	8.544	.790	.430	1.000
Gardening-Tactical	-14.667	8.544	-1.717	.086	1.000
Gardening-Hockey	-25.667	8.544	-3.004	.003	.040
Running-Tactical	-7.917	8.544	-.927	.354	1.000
Running-Hockey	-18.917	8.544	-2.214	.027	.402
Tactical-Hockey	-11.000	8.544	-1.287	.198	1.000

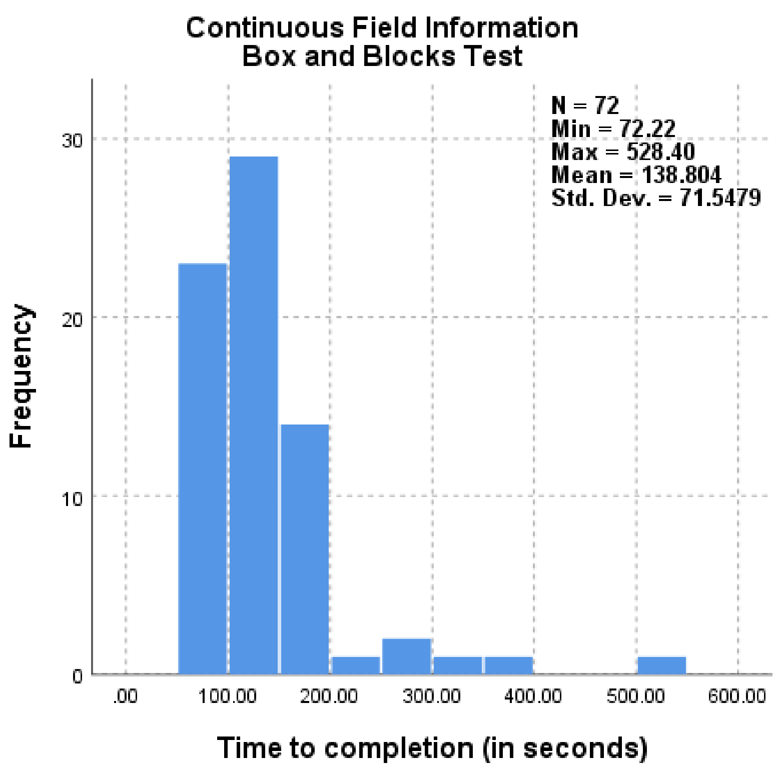
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



Each node shows the sample average rank of Glove.



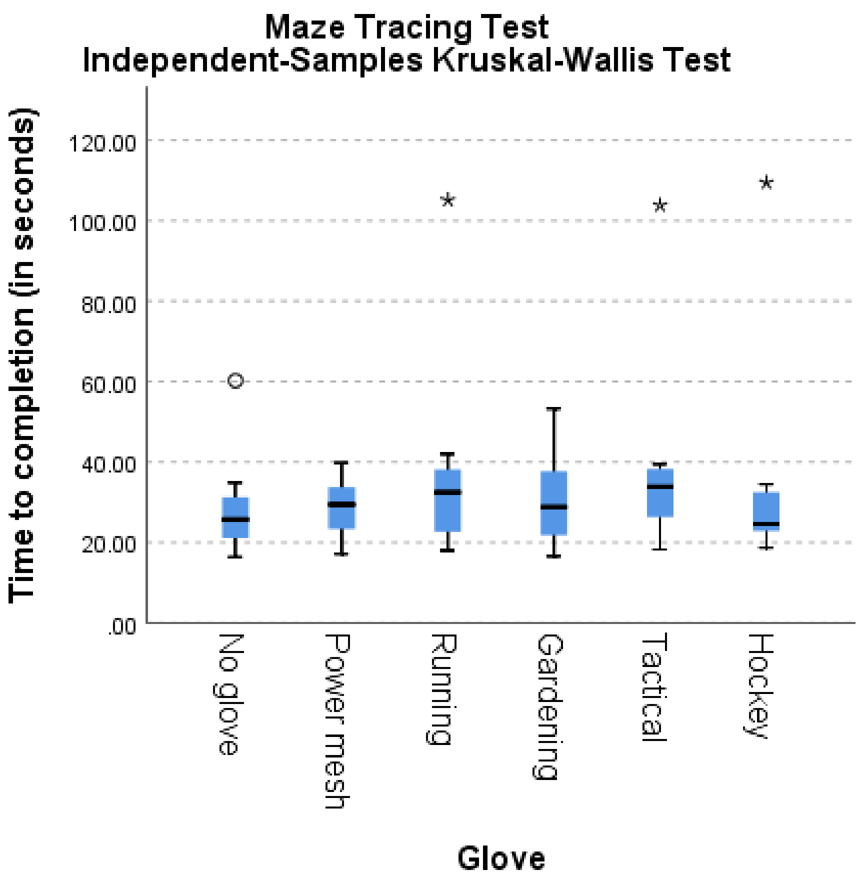
B.1.6 Maze Tracing test

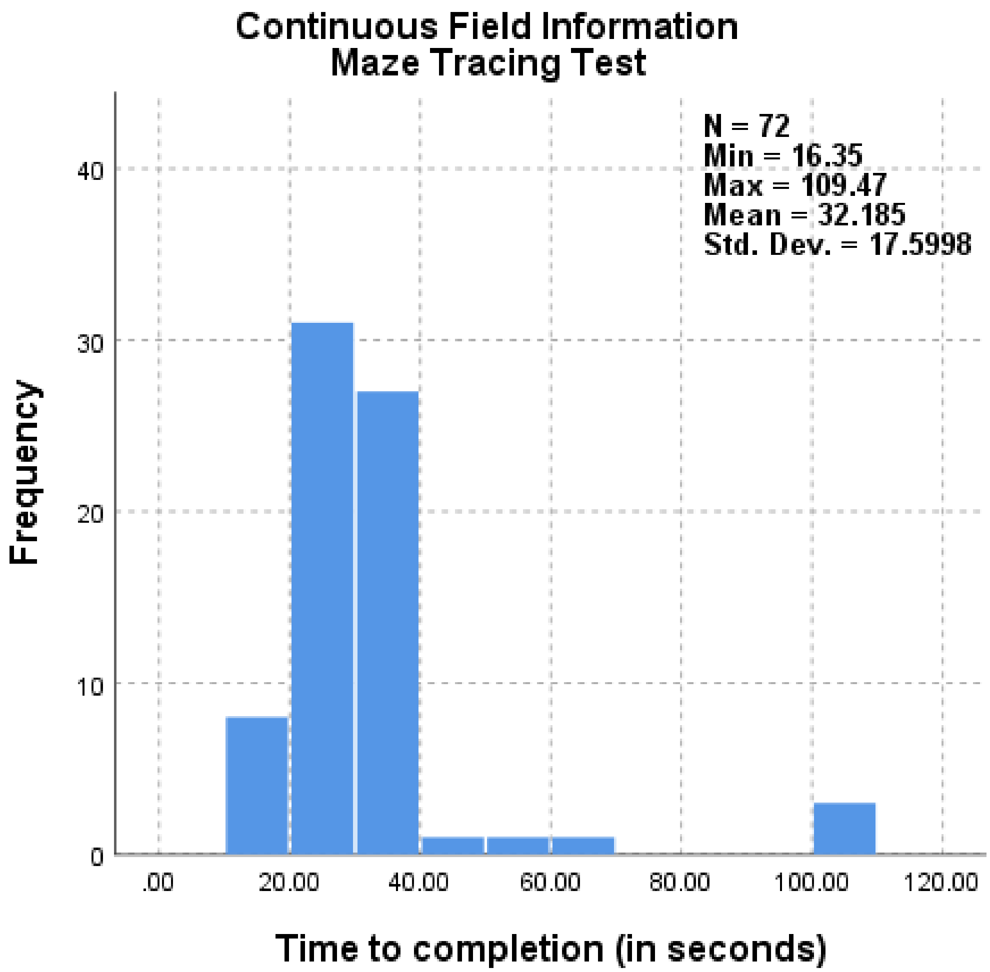
Independent-Samples Kruskal-Wallis Test Summary

Total N	72
Test Statistic	4.493[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.481

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.



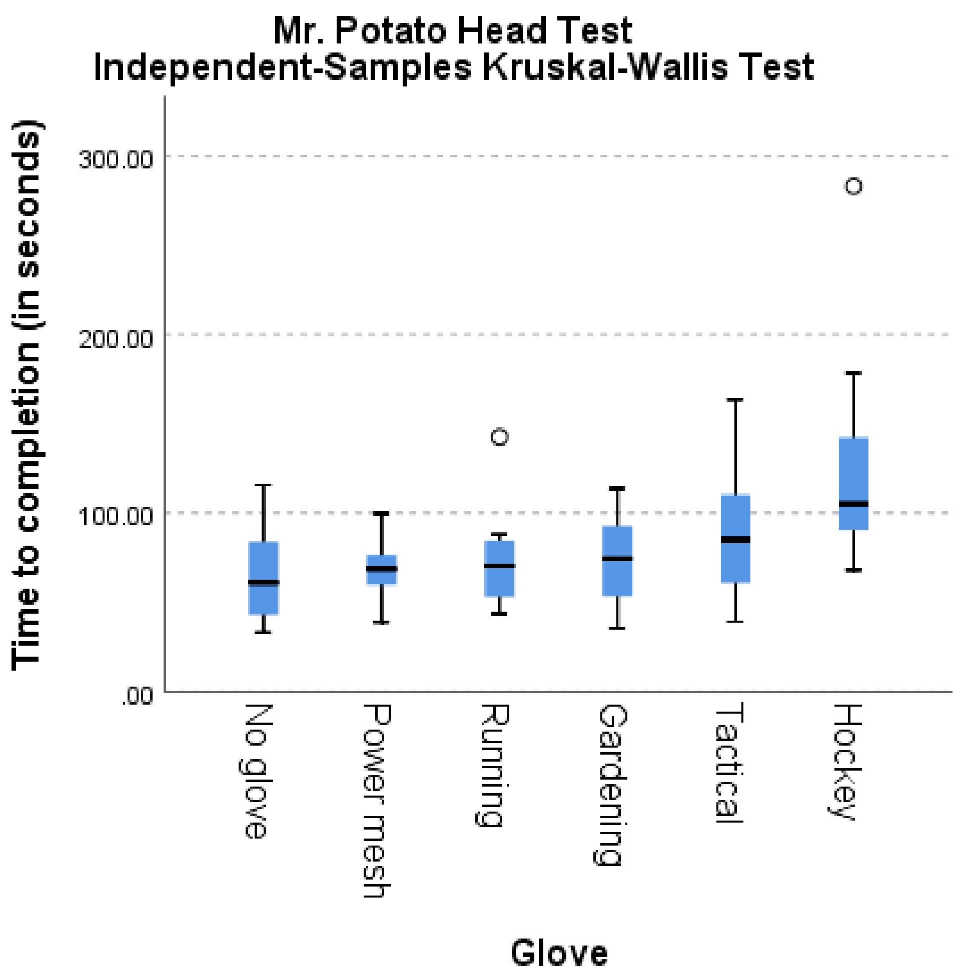

Continuous Field Information
Maze Tracing Test
N = 72
Min = 16.35
Max = 109.47
Mean = 32.185
Std. Dev. = 17.5998
Frequency
0
10
20
30
40
.00
20.00
40.00
60.00
80.00
100.00
120.00
Time to completion (in seconds)

B.1.7 Mr. Potato Head

Independent-Samples Kruskal-Wallis Test Summary

Total N	72
Test Statistic	17.360[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.004

a. The test statistic is adjusted for ties.



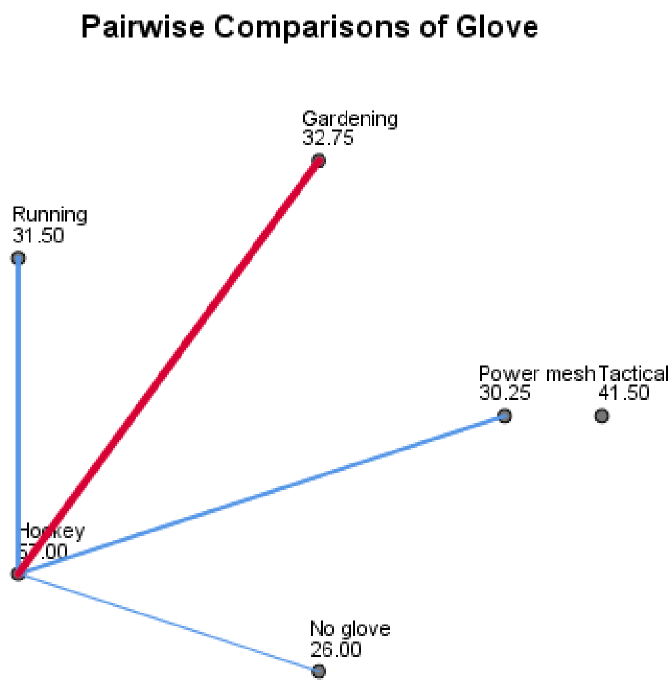
Pairwise Comparisons of Glove

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
No glove-Power mesh	-4.250	8.544	-.497	.619	1.000
No glove-Running	-5.500	8.544	-.644	.520	1.000
No glove-Gardening	-6.750	8.544	-.790	.430	1.000
No glove-Tactical	-15.500	8.544	-1.814	.070	1.000
No glove-Hockey	-31.000	8.544	-3.628	.000	.004
Power mesh-Running	-1.250	8.544	-.146	.884	1.000
Power mesh-Gardening	-2.500	8.544	-.293	.770	1.000
Power mesh-Tactical	-11.250	8.544	-1.317	.188	1.000
Power mesh-Hockey	-26.750	8.544	-3.131	.002	.026
Running-Gardening	-1.250	8.544	-.146	.884	1.000
Running-Tactical	-10.000	8.544	-1.170	.242	1.000
Running-Hockey	-25.500	8.544	-2.985	.003	.043
Gardening-Tactical	-8.750	8.544	-1.024	.306	1.000
Gardening-Hockey	-24.250	8.544	-2.838	.005	.068
Tactical-Hockey	-15.500	8.544	-1.814	.070	1.000

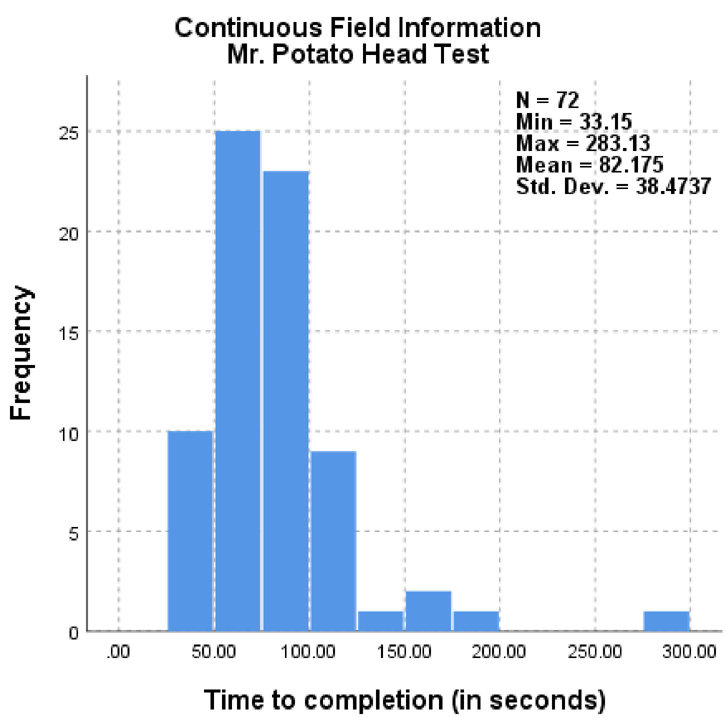
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



Each node shows the sample average rank of Glove.

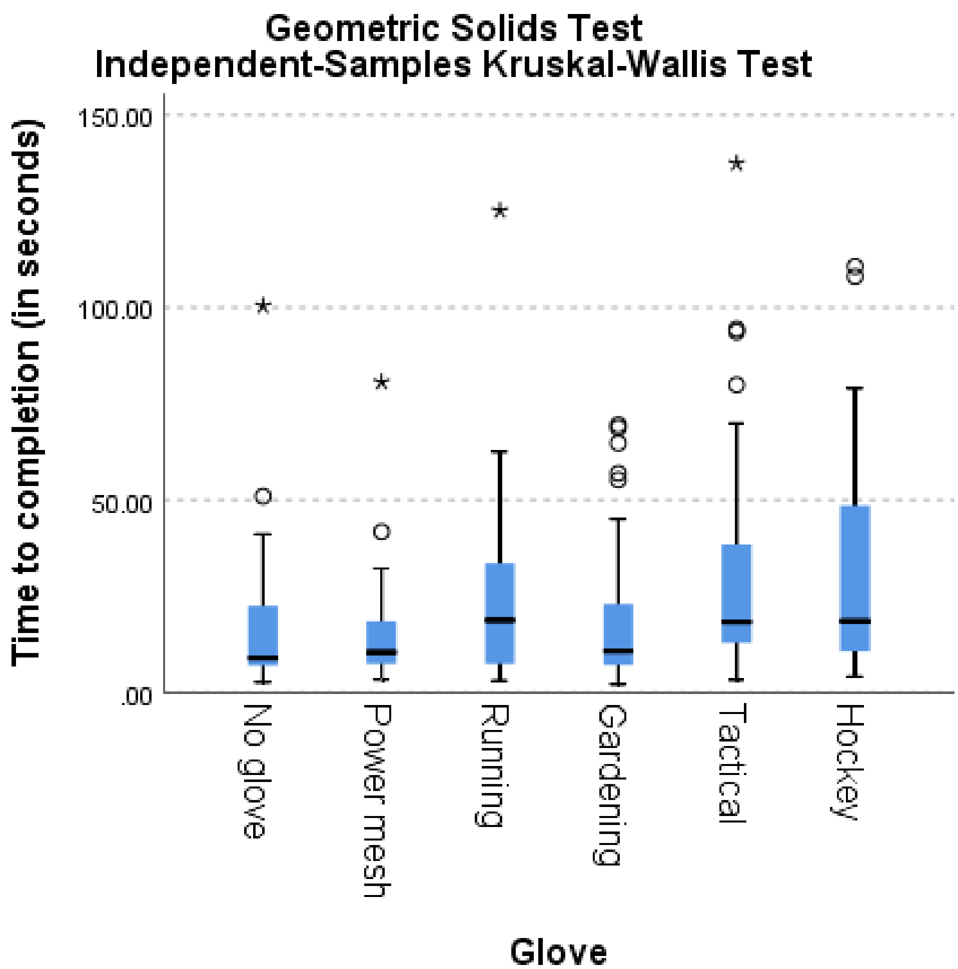


B.1.8 Geometric Solids test

Independent-Samples Kruskal-Wallis Test Summary

Total N	215
Test Statistic	16.610[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.005

a. The test statistic is adjusted for ties.



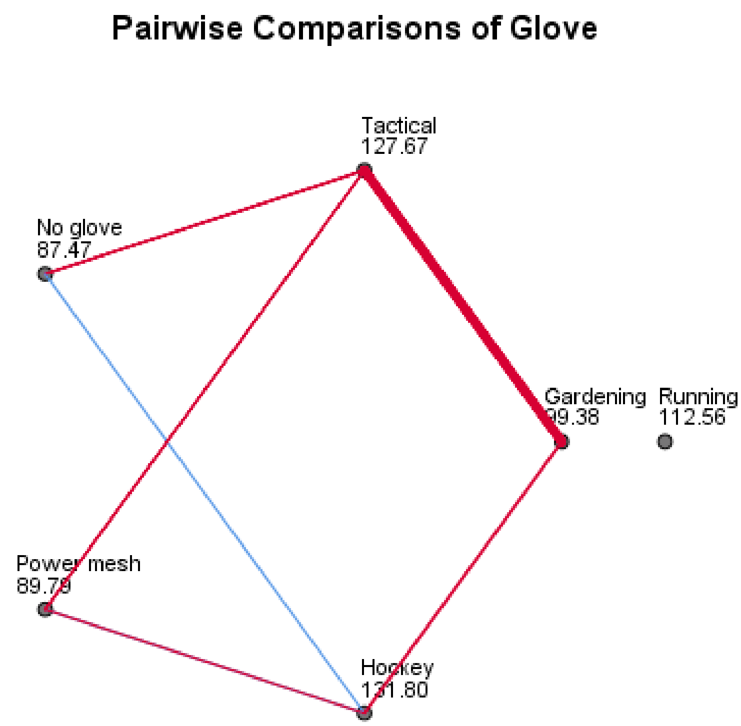
Pairwise Comparisons of Glove

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
No glove-Power mesh	-2.319	14.663	-.158	.874	1.000
No glove-Gardening	-11.903	14.663	-.812	.417	1.000
No glove-Running	-25.083	14.663	-1.711	.087	1.000
No glove-Tactical	-40.194	14.663	-2.741	.006	.092
No glove-Hockey	-44.328	14.767	-3.002	.003	.040
Power mesh-Gardening	-9.583	14.663	-.654	.513	1.000
Power mesh-Running	-22.764	14.663	-1.552	.121	1.000
Power mesh-Tactical	-37.875	14.663	-2.583	.010	.147
Power mesh-Hockey	-42.008	14.767	-2.845	.004	.067
Gardening-Running	13.181	14.663	.899	.369	1.000
Gardening-Tactical	-28.292	14.663	-1.929	.054	.805
Gardening-Hockey	-32.425	14.767	-2.196	.028	.422
Running-Tactical	-15.111	14.663	-1.031	.303	1.000
Running-Hockey	-19.244	14.767	-1.303	.193	1.000
Tactical-Hockey	-4.133	14.767	-.280	.780	1.000

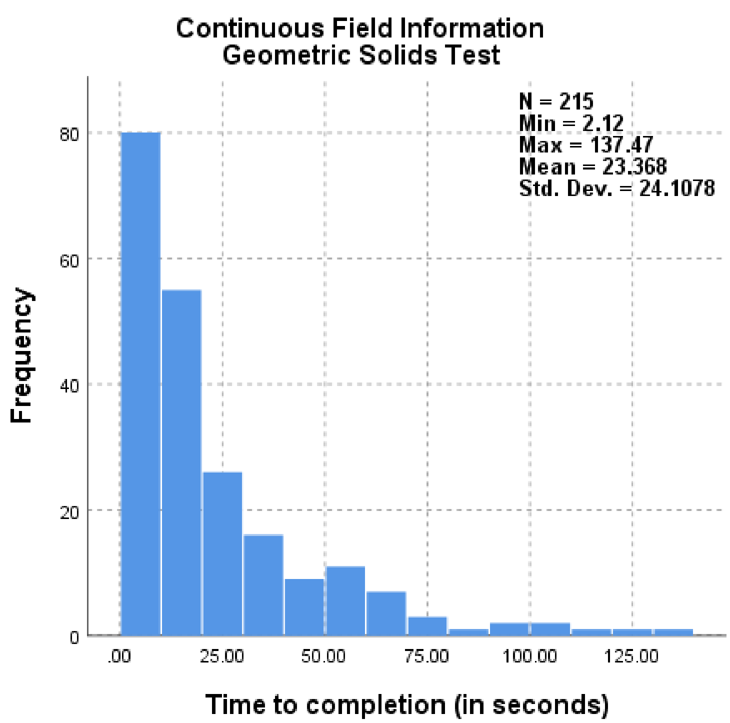
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



Each node shows the sample average rank of Glove.



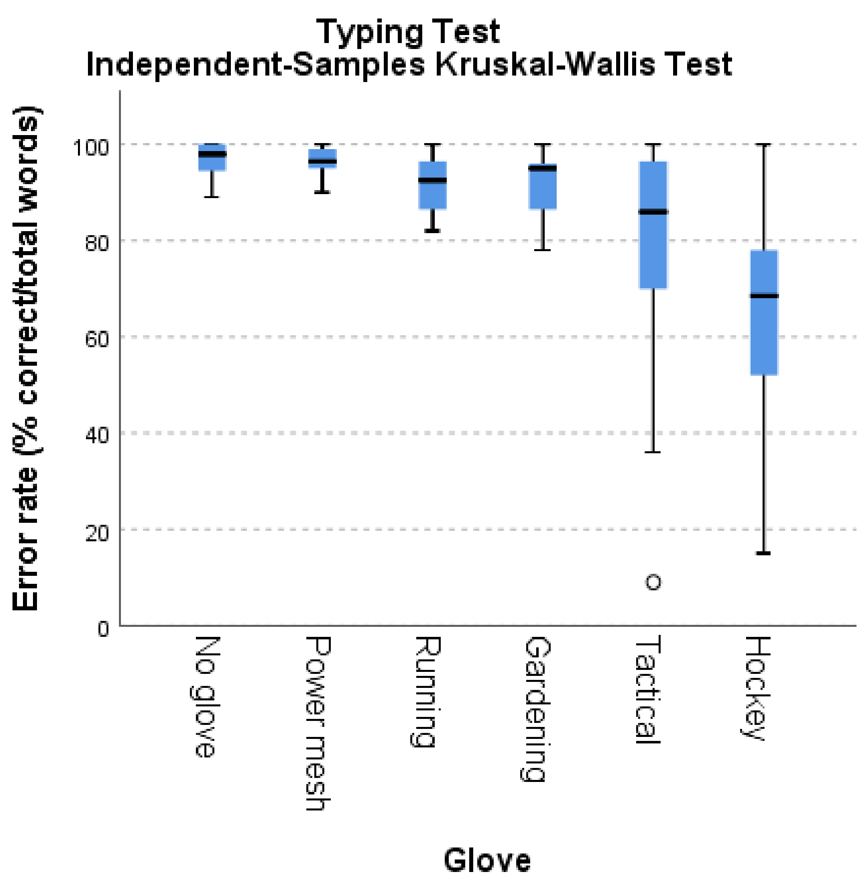
B.2 Error rate data

B.2.1 Typing test – Error rate

Independent-Samples Kruskal-Wallis Test Summary

Total N	72
Test Statistic	27.116[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

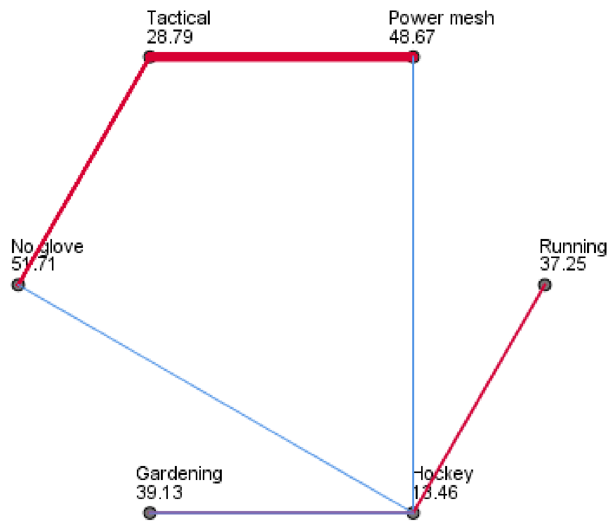
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Hockey-Tactical	15.333	8.489	1.806	.071	1.000
Hockey-Running	23.792	8.489	2.803	.005	.076
Hockey-Gardening	25.667	8.489	3.023	.002	.037
Hockey-Power mesh	35.208	8.489	4.147	.000	.001
Hockey-No glove	38.250	8.489	4.506	.000	.000
Tactical-Running	8.458	8.489	.996	.319	1.000
Tactical-Gardening	10.333	8.489	1.217	.224	1.000
Tactical-Power mesh	19.875	8.489	2.341	.019	.288
Tactical-No glove	22.917	8.489	2.699	.007	.104
Running-Gardening	-1.875	8.489	-.221	.825	1.000
Running-Power mesh	11.417	8.489	1.345	.179	1.000
Running-No glove	14.458	8.489	1.703	.089	1.000
Gardening-Power mesh	9.542	8.489	1.124	.261	1.000
Gardening-No glove	12.583	8.489	1.482	.138	1.000
Power mesh-No glove	3.042	8.489	.358	.720	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

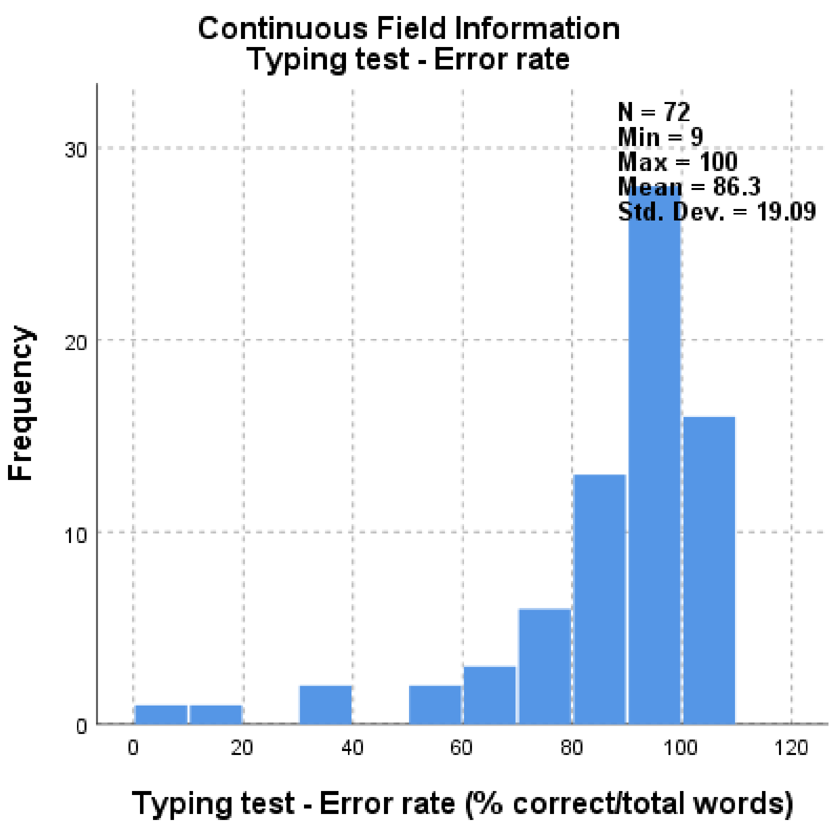
Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.



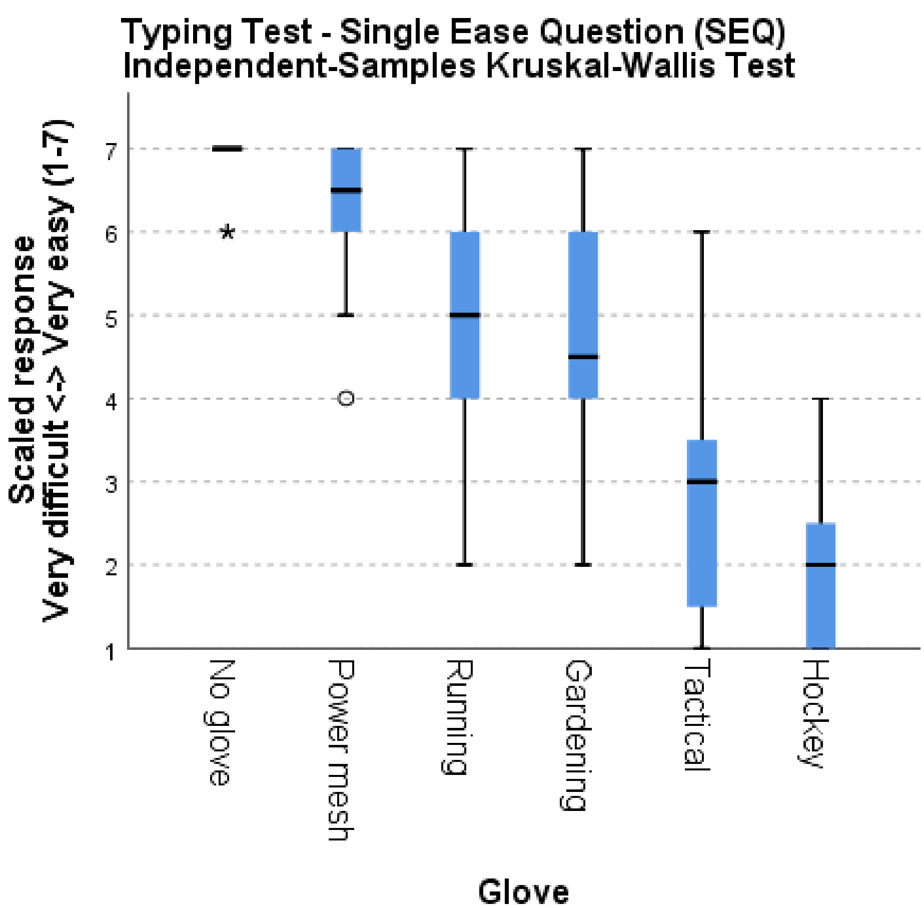
B.3 Single Ease Question (SEQ) data

B.3.1 Typing test – SEQ score

Independent-Samples Kruskal-Wallis Test Summary

Total N	72
Test Statistic	50.383[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

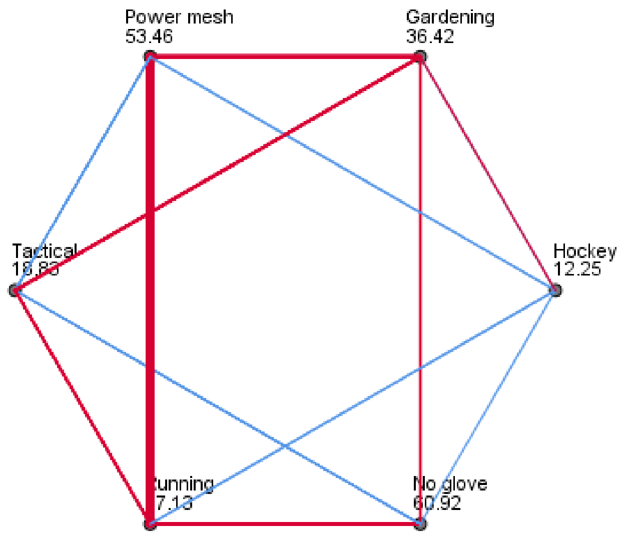
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Hockey-Tactical	6.583	8.416	.782	.434	1.000
Hockey-Gardening	24.167	8.416	2.871	.004	.061
Hockey-Running	24.875	8.416	2.956	.003	.047
Hockey-Power mesh	41.208	8.416	4.896	.000	.000
Hockey-No glove	48.667	8.416	5.783	.000	.000
Tactical-Gardening	17.583	8.416	2.089	.037	.550
Tactical-Running	18.292	8.416	2.173	.030	.446
Tactical-Power mesh	34.625	8.416	4.114	.000	.001
Tactical-No glove	42.083	8.416	5.000	.000	.000
Gardening-Running	.708	8.416	.084	.933	1.000
Gardening-Power mesh	17.042	8.416	2.025	.043	.643
Gardening-No glove	24.500	8.416	2.911	.004	.054
Running-Power mesh	16.333	8.416	1.941	.052	.784
Running-No glove	23.792	8.416	2.827	.005	.070
Power mesh-No glove	7.458	8.416	.886	.376	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

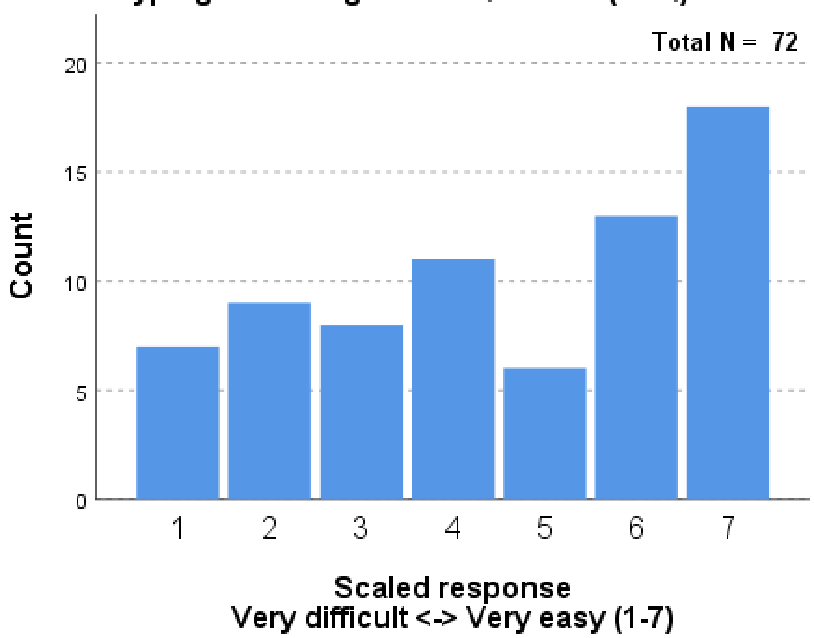
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.

Categorical Field Information
Typing test - Single Ease Question (SEQ)



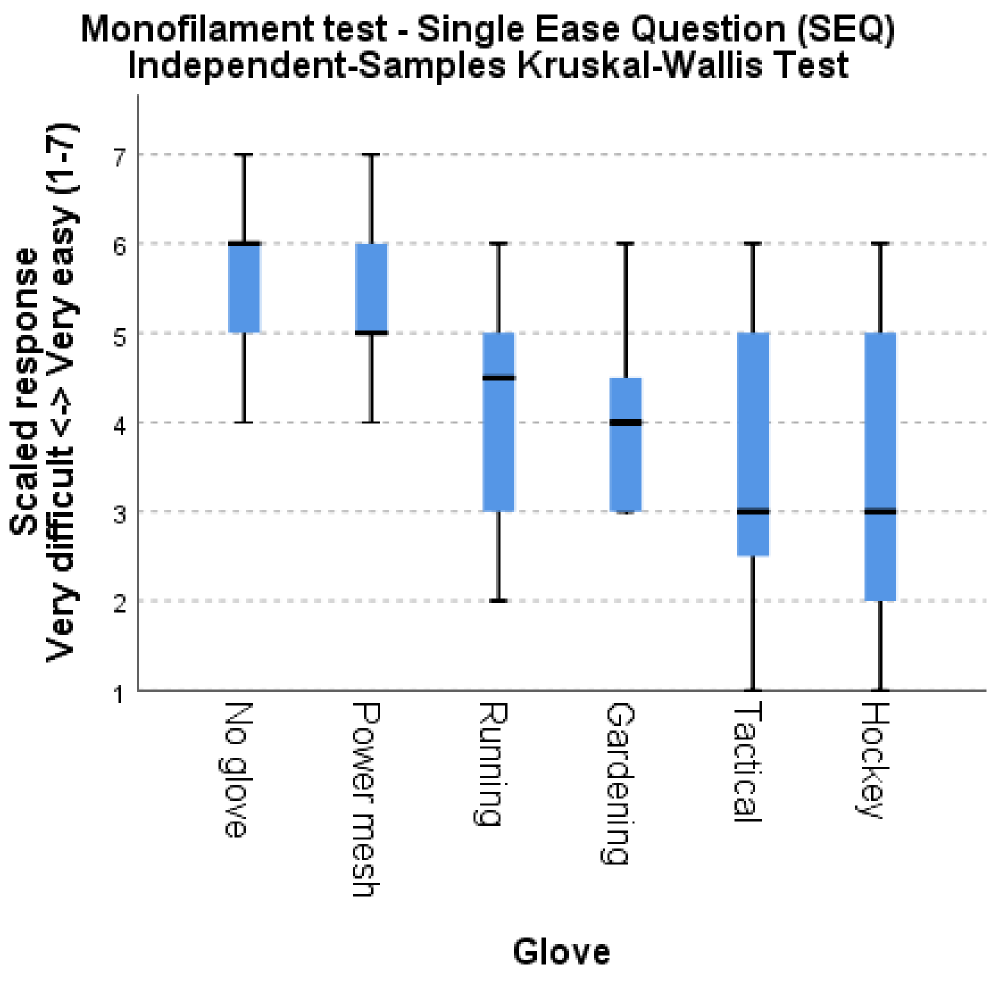
Typing test
(answer scaled response questions on the next page) - Single Ease Question (1-7) field is ordinal but is treated as continuous in the test.

B.3.2 Monofilament test – SEQ score

Independent-Samples Kruskal-Wallis Test Summary

Total N	72
Test Statistic	25.947[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



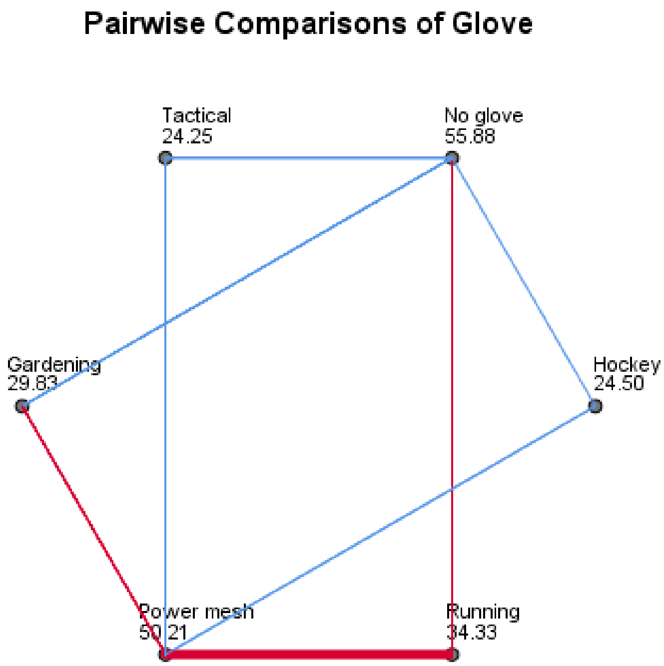
Pairwise Comparisons of Glove

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Tactical-Hockey	-.250	8.359	-.030	.976	1.000
Tactical-Gardening	5.583	8.359	.668	.504	1.000
Tactical-Running	10.083	8.359	1.206	.228	1.000
Tactical-Power mesh	25.958	8.359	3.105	.002	.029
Tactical-No glove	31.625	8.359	3.783	.000	.002
Hockey-Gardening	5.333	8.359	.638	.523	1.000
Hockey-Running	9.833	8.359	1.176	.239	1.000
Hockey-Power mesh	25.708	8.359	3.076	.002	.032
Hockey-No glove	31.375	8.359	3.753	.000	.003
Gardening-Running	4.500	8.359	.538	.590	1.000
Gardening-Power mesh	20.375	8.359	2.437	.015	.222
Gardening-No glove	26.042	8.359	3.115	.002	.028
Running-Power mesh	15.875	8.359	1.899	.058	.863
Running-No glove	21.542	8.359	2.577	.010	.149
Power mesh-No glove	5.667	8.359	.678	.498	1.000

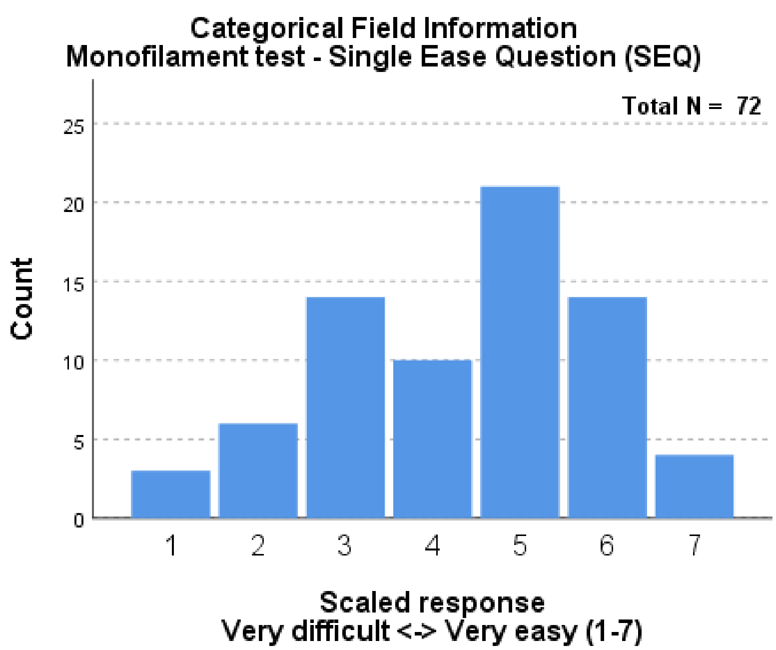
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



Each node shows the sample average rank of Glove.



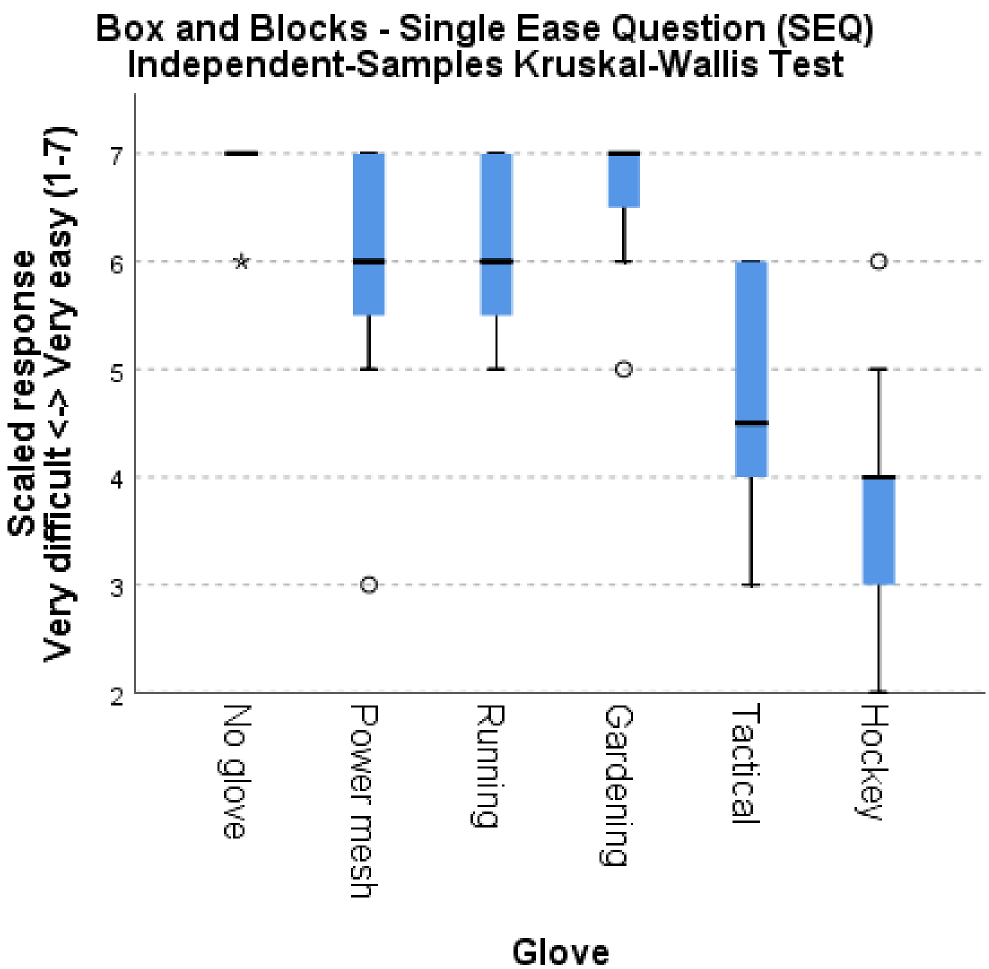
Monofilament test

Please enter the smallest force registered on each finger (in g) - Single Ease Question (1-7) field is ordinal but is treated as continuous in the test.

B.3.3 Box and Blocks – SEQ score

Independent-Samples Kruskal-Wallis Test Summary	
Total N	72
Test Statistic	44.197[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

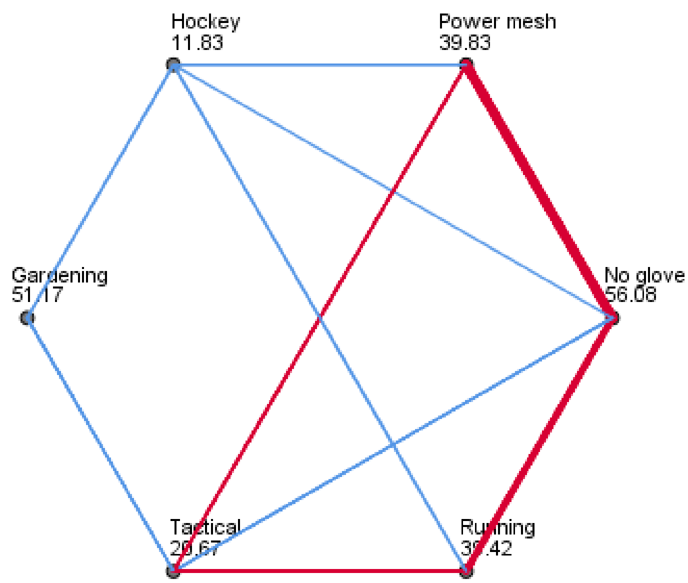
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Hockey-Tactical	8.833	8.176	1.080	.280	1.000
Hockey-Running	27.583	8.176	3.374	.001	.011
Hockey-Power mesh	28.000	8.176	3.424	.001	.009
Hockey-Gardening	39.333	8.176	4.811	.000	.000
Hockey-No glove	44.250	8.176	5.412	.000	.000
Tactical-Running	18.750	8.176	2.293	.022	.328
Tactical-Power mesh	19.167	8.176	2.344	.019	.286
Tactical-Gardening	30.500	8.176	3.730	.000	.003
Tactical-No glove	35.417	8.176	4.332	.000	.000
Running-Power mesh	.417	8.176	.051	.959	1.000
Running-Gardening	-11.750	8.176	-1.437	.151	1.000
Running-No glove	16.667	8.176	2.038	.042	.623
Power mesh-Gardening	-11.333	8.176	-1.386	.166	1.000
Power mesh-No glove	16.250	8.176	1.987	.047	.703
Gardening-No glove	4.917	8.176	.601	.548	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

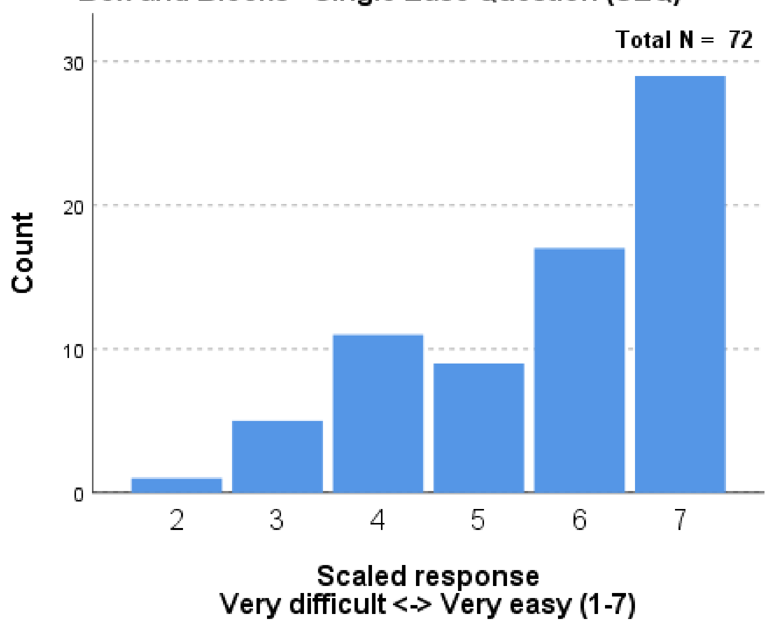
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.

Categorical Field Information
Box and Blocks - Single Ease Question (SEQ)



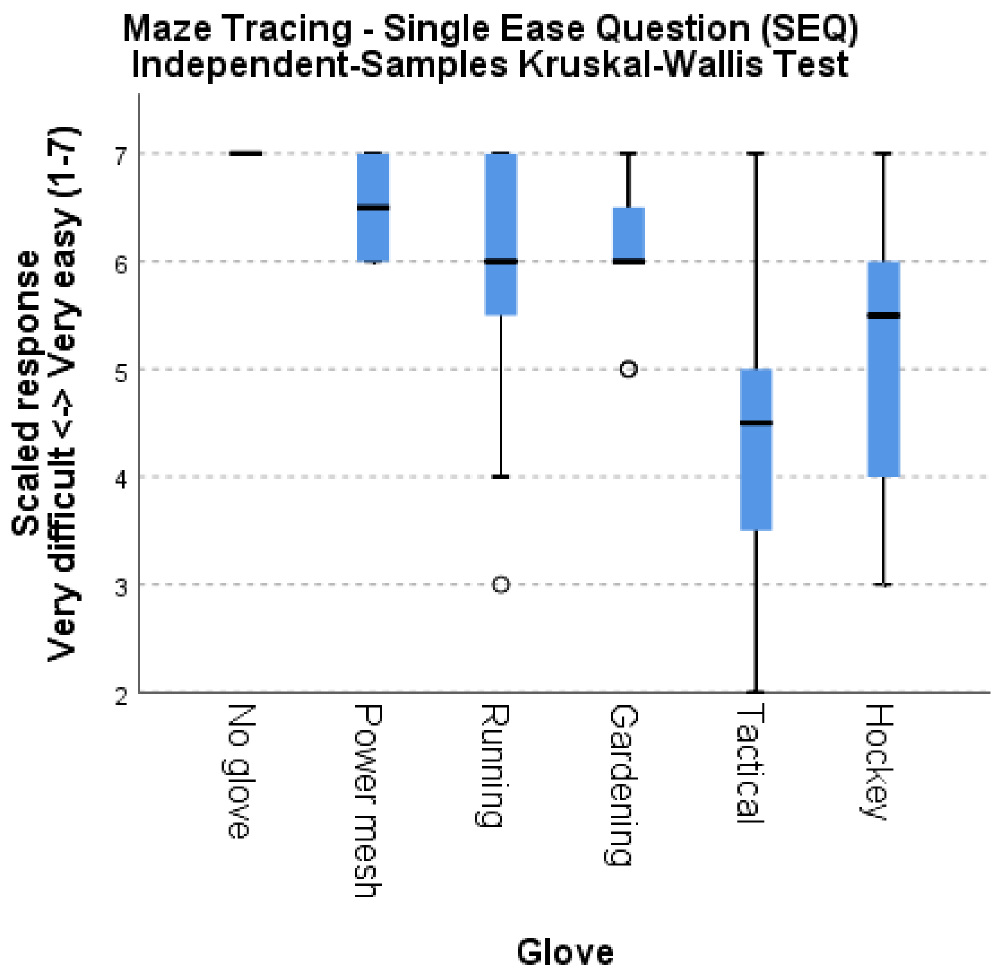
Box and Blocks - Single Ease Question (1-7) field is ordinal but is treated as continuous in the test.

B.3.4 Maze Tracing – SEQ score

Independent-Samples Kruskal-Wallis Test Summary

Total N	72
Test Statistic	32.960[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

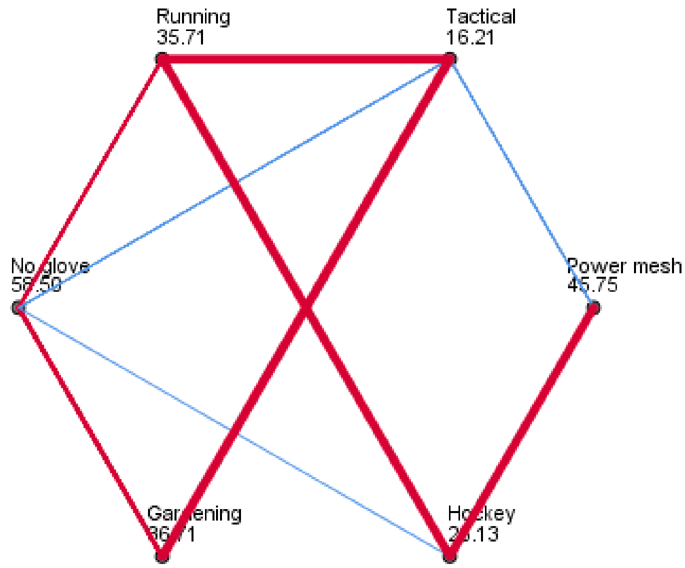
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Tactical-Hockey	-9.917	8.131	-1.220	.223	1.000
Tactical-Running	19.500	8.131	2.398	.016	.247
Tactical-Gardening	20.500	8.131	2.521	.012	.175
Tactical-Power mesh	29.542	8.131	3.633	.000	.004
Tactical-No glove	42.292	8.131	5.201	.000	.000
Hockey-Running	9.583	8.131	1.179	.239	1.000
Hockey-Gardening	10.583	8.131	1.302	.193	1.000
Hockey-Power mesh	19.625	8.131	2.414	.016	.237
Hockey-No glove	32.375	8.131	3.982	.000	.001
Running-Gardening	-1.000	8.131	-.123	.902	1.000
Running-Power mesh	10.042	8.131	1.235	.217	1.000
Running-No glove	22.792	8.131	2.803	.005	.076
Gardening-Power mesh	9.042	8.131	1.112	.266	1.000
Gardening-No glove	21.792	8.131	2.680	.007	.110
Power mesh-No glove	12.750	8.131	1.568	.117	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

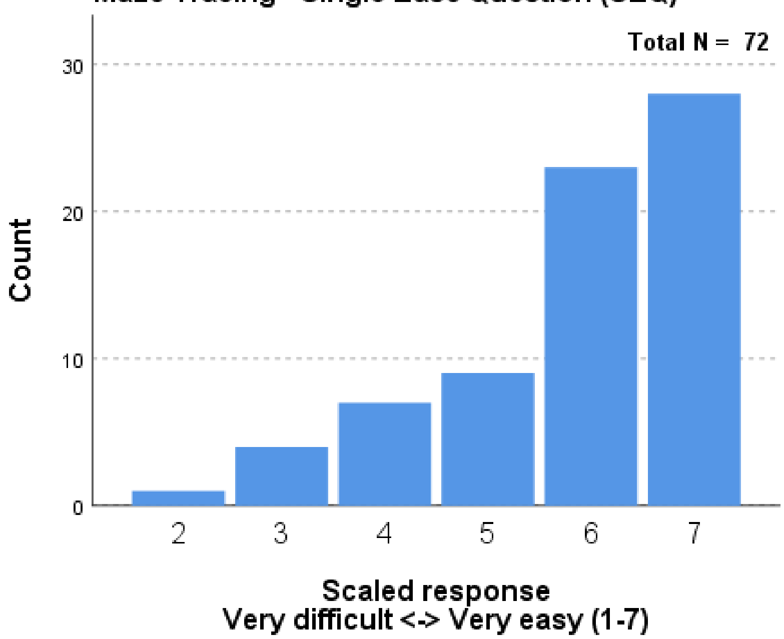
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.

Categorical Field Information
Maze Tracing - Single Ease Question (SEQ)



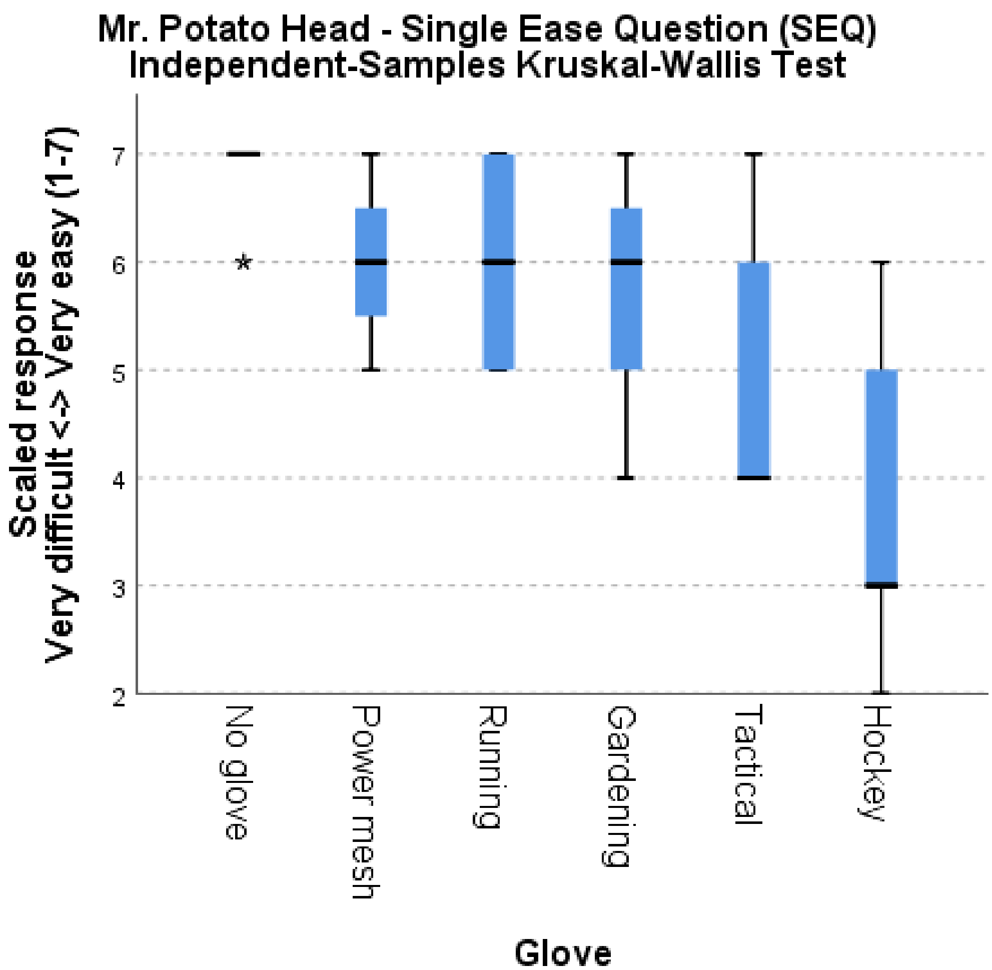
Maze Tracing - Single Ease Question (1-7) field is ordinal but is treated as continuous in the test.

B.3.5 Mr. Potato Head – SEQ score

Independent-Samples Kruskal-Wallis Test Summary

Total N	72
Test Statistic	34.049[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

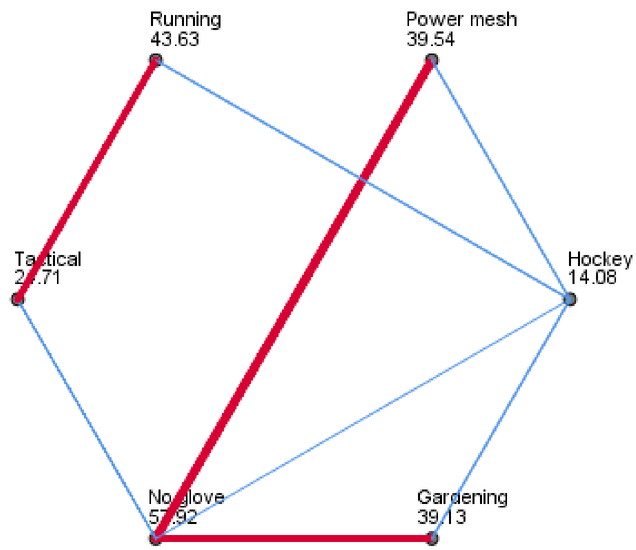
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Hockey-Tactical	10.625	8.280	1.283	.199	1.000
Hockey-Gardening	25.042	8.280	3.024	.002	.037
Hockey-Power mesh	25.458	8.280	3.075	.002	.032
Hockey-Running	29.542	8.280	3.568	.000	.005
Hockey-No glove	43.833	8.280	5.294	.000	.000
Tactical-Gardening	14.417	8.280	1.741	.082	1.000
Tactical-Power mesh	14.833	8.280	1.791	.073	1.000
Tactical-Running	18.917	8.280	2.285	.022	.335
Tactical-No glove	33.208	8.280	4.011	.000	.001
Gardening-Power mesh	.417	8.280	.050	.960	1.000
Gardening-Running	4.500	8.280	.543	.587	1.000
Gardening-No glove	18.792	8.280	2.270	.023	.349
Power mesh-Running	-4.083	8.280	-.493	.622	1.000
Power mesh-No glove	18.375	8.280	2.219	.026	.397
Running-No glove	14.292	8.280	1.726	.084	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

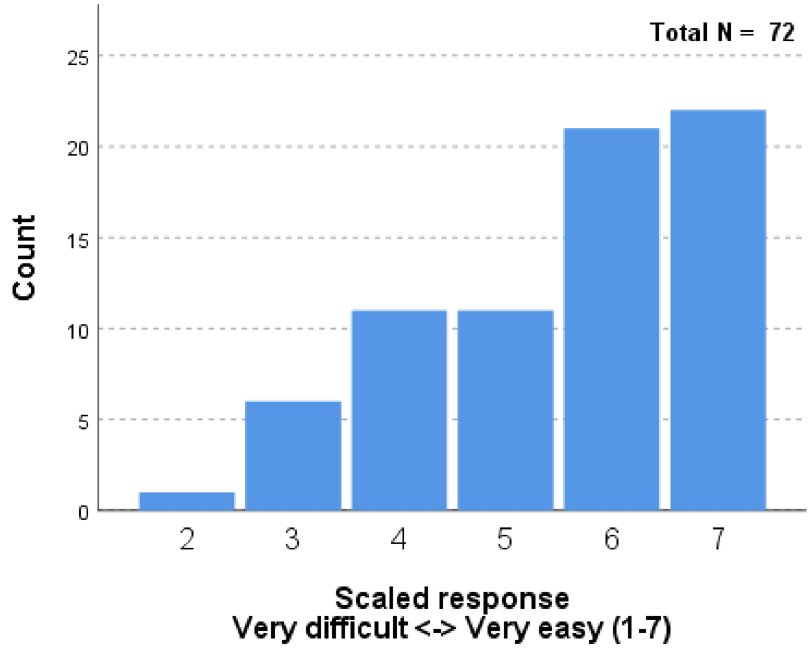
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.

Categorical Field Information
Mr. Potato Head - Single Ease Question (SEQ)



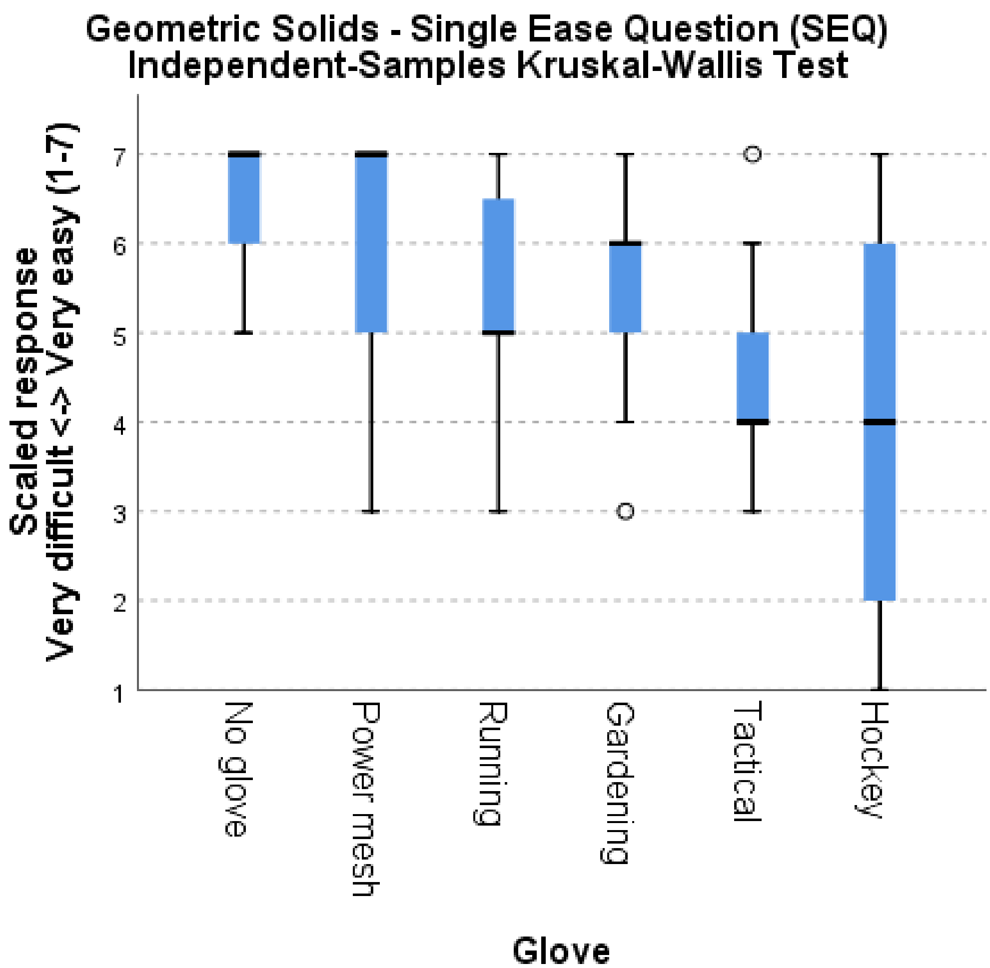
Mr. Potato Head - Single Ease Question (1-7) field is ordinal but is treated as continuous in the test.

B.3.6 Geometric Solids - Single Ease Question (1-7) across Glove

Independent-Samples Kruskal-Wallis Test Summary

Total N	72
Test Statistic	19.294[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.002

a. The test statistic is adjusted for ties.



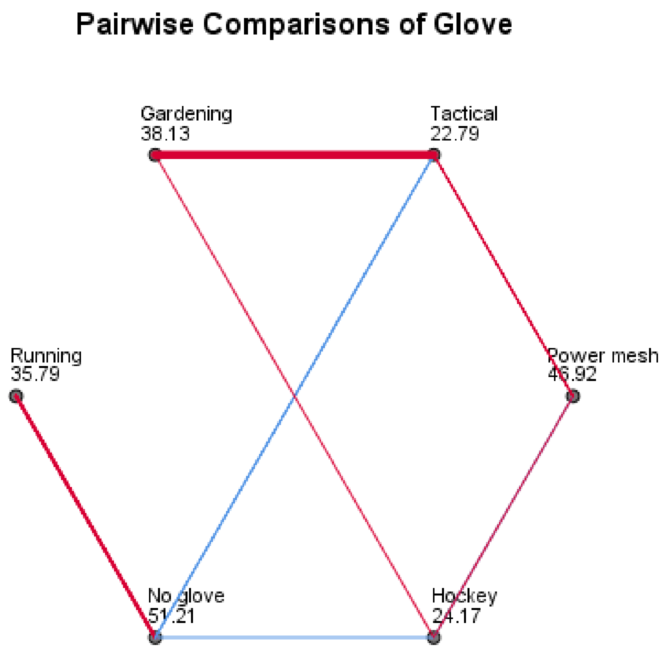
Pairwise Comparisons of Glove

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Tactical-Hockey	-1.375	8.321	-.165	.869	1.000
Tactical-Running	13.000	8.321	1.562	.118	1.000
Tactical-Gardening	15.333	8.321	1.843	.065	.981
Tactical-Power mesh	24.125	8.321	2.899	.004	.056
Tactical-No glove	28.417	8.321	3.415	.001	.010
Hockey-Running	11.625	8.321	1.397	.162	1.000
Hockey-Gardening	13.958	8.321	1.677	.093	1.000
Hockey-Power mesh	22.750	8.321	2.734	.006	.094
Hockey-No glove	27.042	8.321	3.250	.001	.017
Running-Gardening	-2.333	8.321	-.280	.779	1.000
Running-Power mesh	11.125	8.321	1.337	.181	1.000
Running-No glove	15.417	8.321	1.853	.064	.959
Gardening-Power mesh	8.792	8.321	1.057	.291	1.000
Gardening-No glove	13.083	8.321	1.572	.116	1.000
Power mesh-No glove	4.292	8.321	.516	.606	1.000

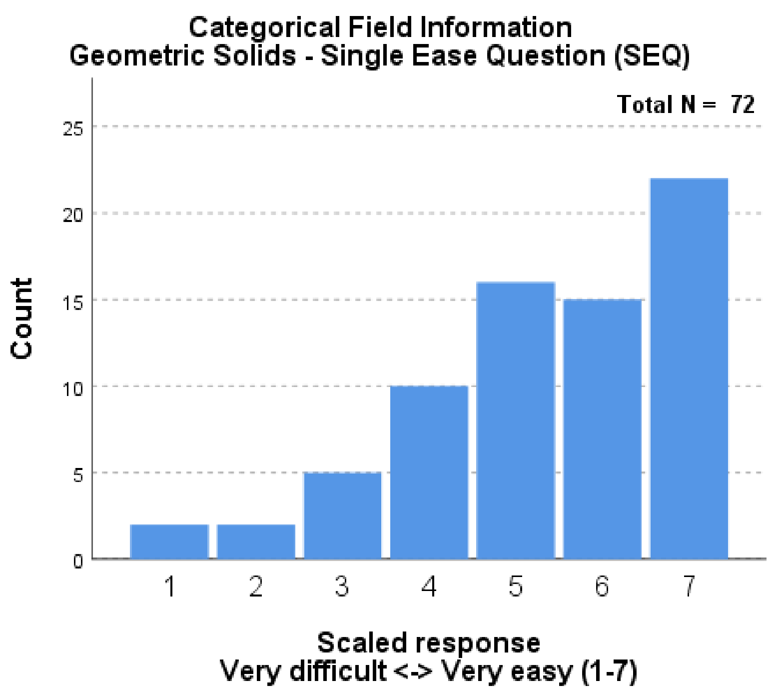
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



Each node shows the sample average rank of Glove.



Geometric Solids - Single Ease Question (1-7) field is ordinal but is treated as continuous in the test.

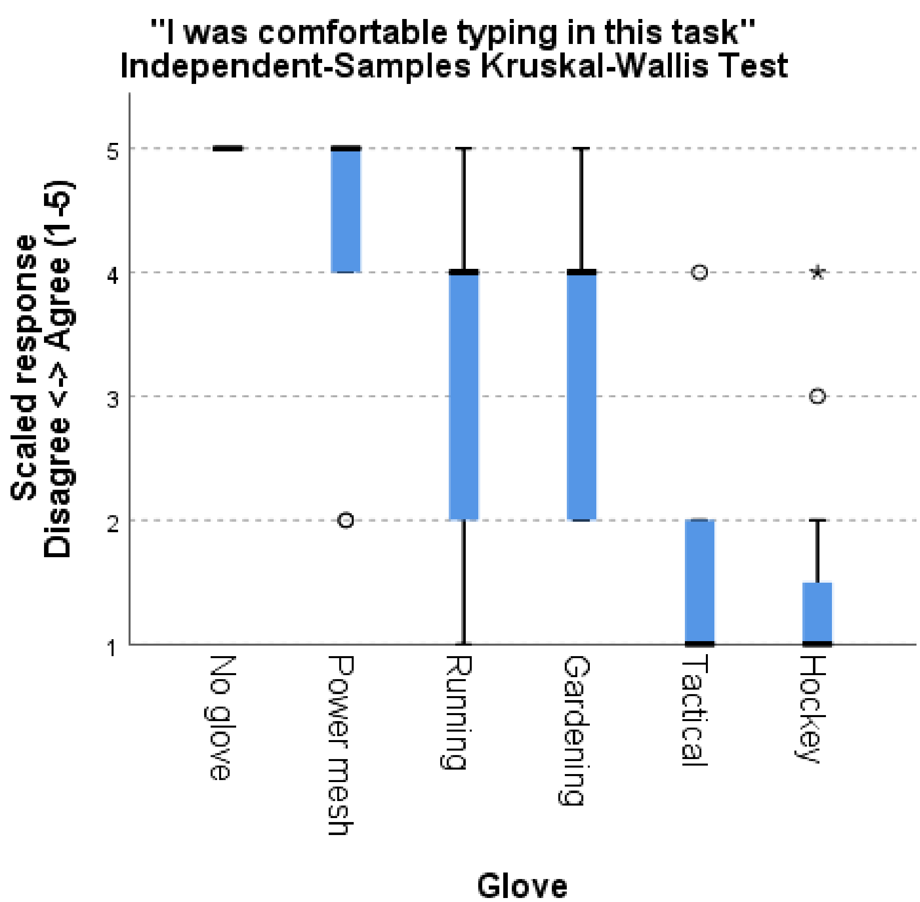
B.4 Scaled response data

B.4.1 "I was comfortable typing in this task"

Independent-Samples Kruskal-Wallis Test Summary

Total N	72
Test Statistic	49.740[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

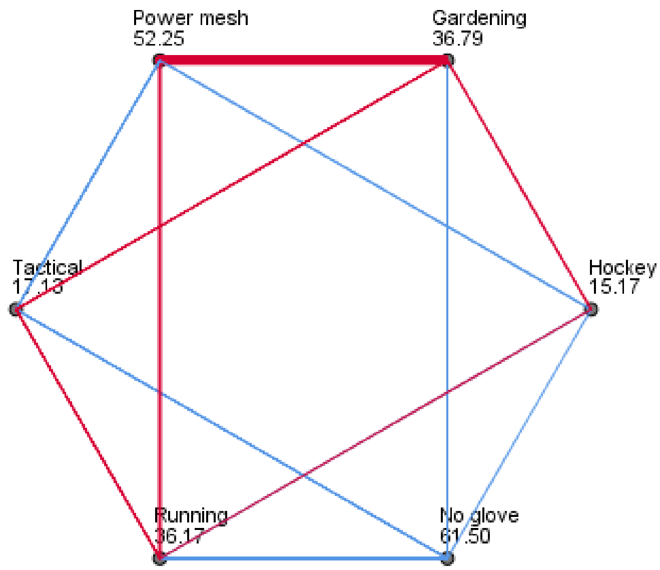
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Hockey-Tactical	1.958	8.277	.237	.813	1.000
Hockey-Running	21.000	8.277	2.537	.011	.168
Hockey-Gardening	21.625	8.277	2.613	.009	.135
Hockey-Power mesh	37.083	8.277	4.480	.000	.000
Hockey-No glove	46.333	8.277	5.598	.000	.000
Tactical-Running	19.042	8.277	2.301	.021	.321
Tactical-Gardening	19.667	8.277	2.376	.017	.262
Tactical-Power mesh	35.125	8.277	4.244	.000	.000
Tactical-No glove	44.375	8.277	5.361	.000	.000
Running-Gardening	-.625	8.277	-.076	.940	1.000
Running-Power mesh	16.083	8.277	1.943	.052	.780
Running-No glove	25.333	8.277	3.061	.002	.033
Gardening-Power mesh	15.458	8.277	1.868	.062	.927
Gardening-No glove	24.708	8.277	2.985	.003	.043
Power mesh-No glove	9.250	8.277	1.118	.264	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

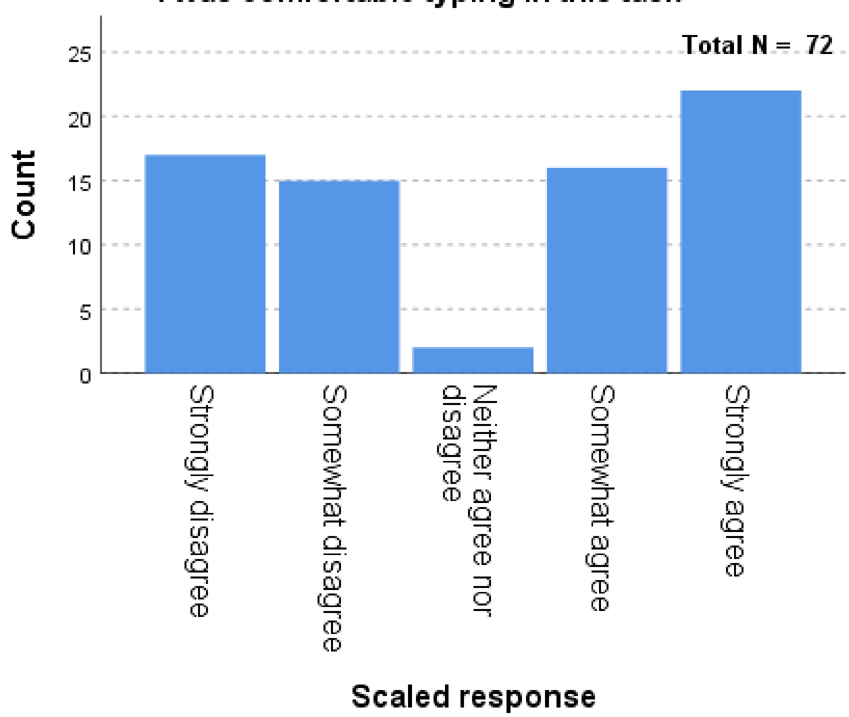
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.

Categorical Field Information
"I was comfortable typing in this task"

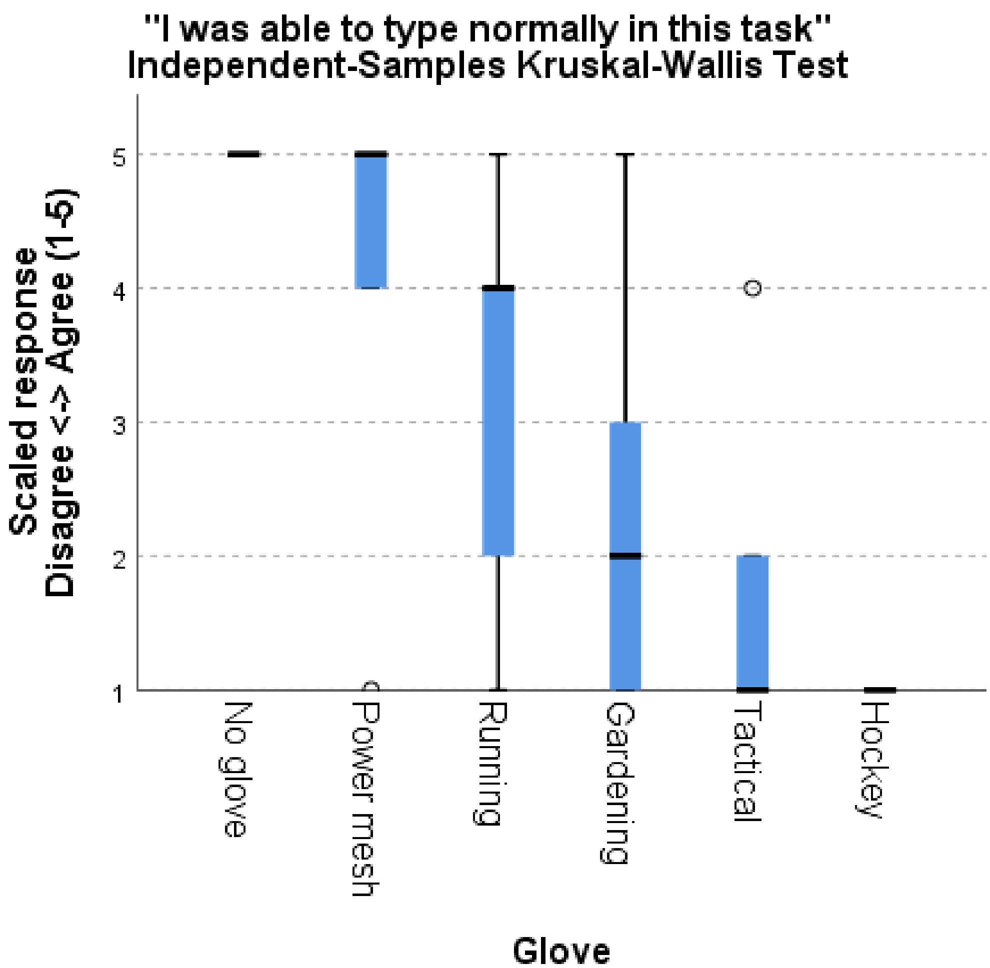


B.4.2 *"I was able to type normally in this task"*

Independent-Samples Kruskal-Wallis Test Summary

Total N	72
Test Statistic	49.606[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

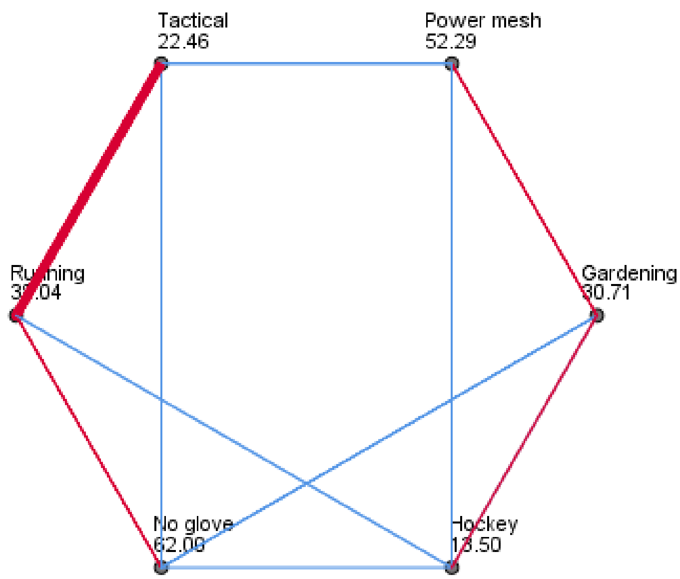
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Hockey-Tactical	8.958	8.185	1.094	.274	1.000
Hockey-Gardening	17.208	8.185	2.102	.036	.533
Hockey-Running	24.542	8.185	2.998	.003	.041
Hockey-Power mesh	38.792	8.185	4.739	.000	.000
Hockey-No glove	48.500	8.185	5.925	.000	.000
Tactical-Gardening	8.250	8.185	1.008	.313	1.000
Tactical-Running	15.583	8.185	1.904	.057	.854
Tactical-Power mesh	29.833	8.185	3.645	.000	.004
Tactical-No glove	39.542	8.185	4.831	.000	.000
Gardening-Running	7.333	8.185	.896	.370	1.000
Gardening-Power mesh	21.583	8.185	2.637	.008	.126
Gardening-No glove	31.292	8.185	3.823	.000	.002
Running-Power mesh	14.250	8.185	1.741	.082	1.000
Running-No glove	23.958	8.185	2.927	.003	.051
Power mesh-No glove	9.708	8.185	1.186	.236	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

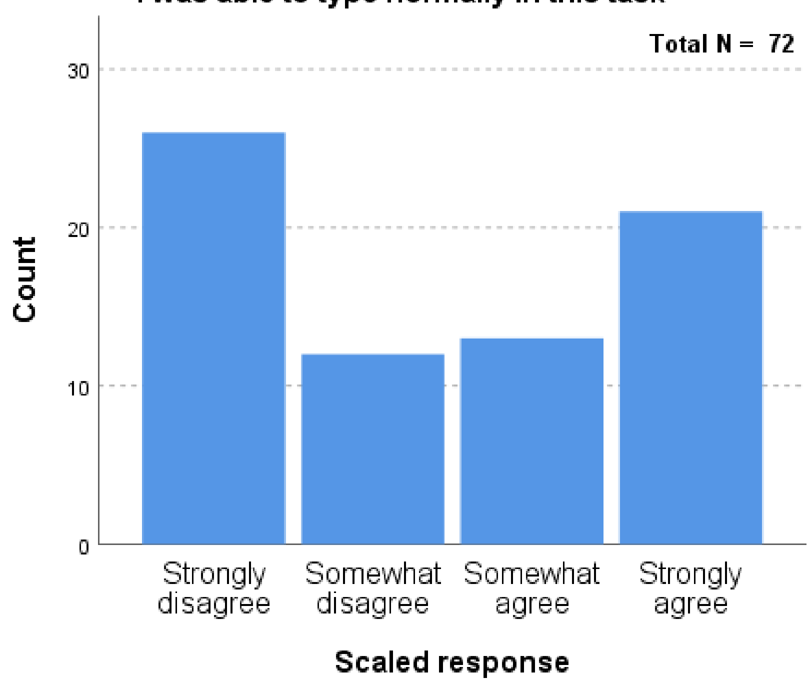
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.

Categorical Field Information
"I was able to type normally in this task"

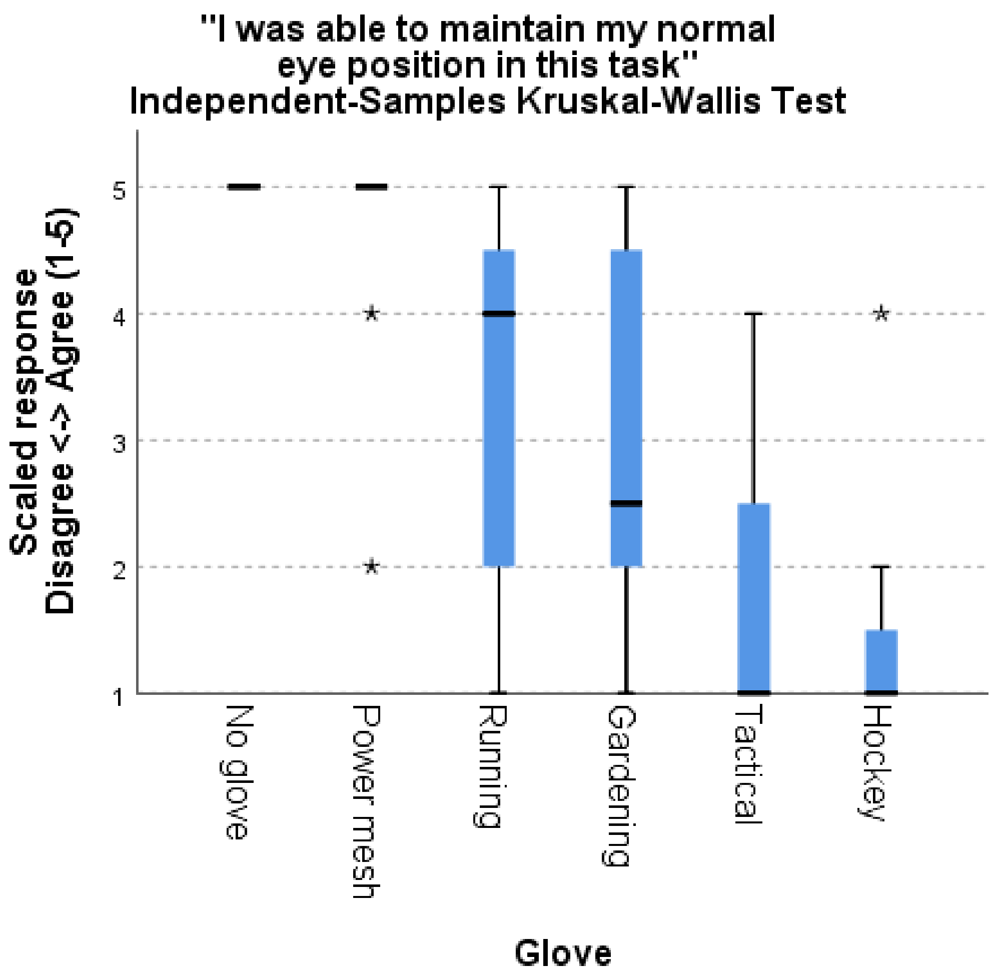


B.4.3 "I was able to maintain my normal eye position in this task"

Independent-Samples Kruskal-Wallis Test Summary

Total N	72
Test Statistic	46.548[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

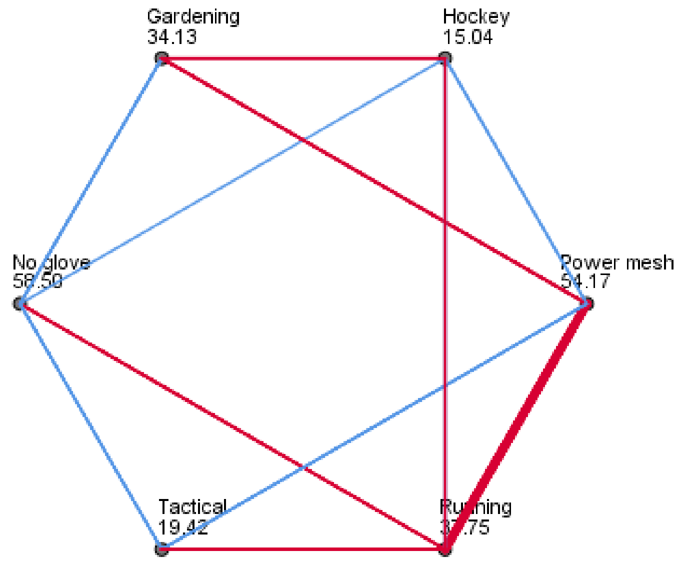
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Hockey-Tactical	4.375	8.176	.535	.593	1.000
Hockey-Gardening	19.083	8.176	2.334	.020	.294
Hockey-Running	22.708	8.176	2.778	.005	.082
Hockey-Power mesh	39.125	8.176	4.786	.000	.000
Hockey-No glove	43.458	8.176	5.316	.000	.000
Tactical-Gardening	14.708	8.176	1.799	.072	1.000
Tactical-Running	18.333	8.176	2.242	.025	.374
Tactical-Power mesh	34.750	8.176	4.250	.000	.000
Tactical-No glove	39.083	8.176	4.781	.000	.000
Gardening-Running	3.625	8.176	.443	.657	1.000
Gardening-Power mesh	20.042	8.176	2.451	.014	.213
Gardening-No glove	24.375	8.176	2.981	.003	.043
Running-Power mesh	16.417	8.176	2.008	.045	.670
Running-No glove	20.750	8.176	2.538	.011	.167
Power mesh-No glove	4.333	8.176	.530	.596	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

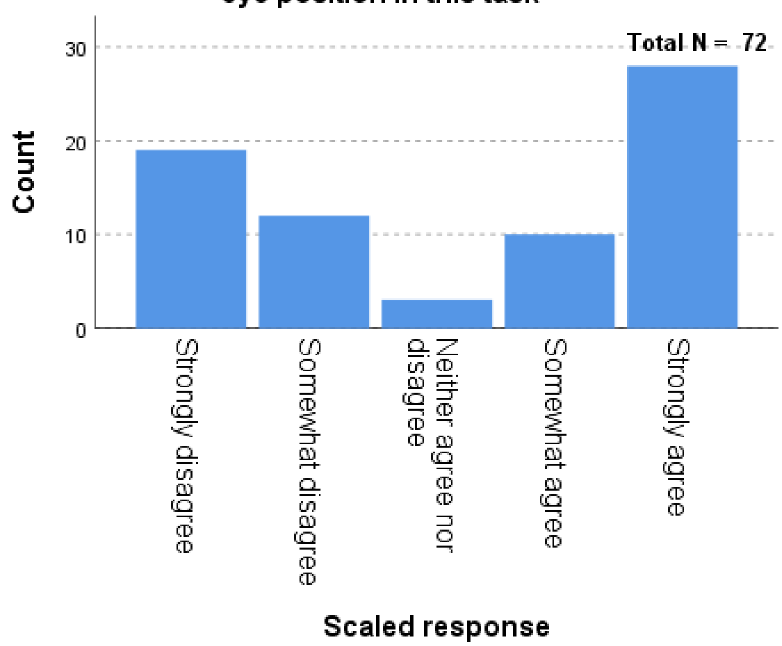
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.

Categorical Field Information
"I was able to maintain my normal eye position in this task"

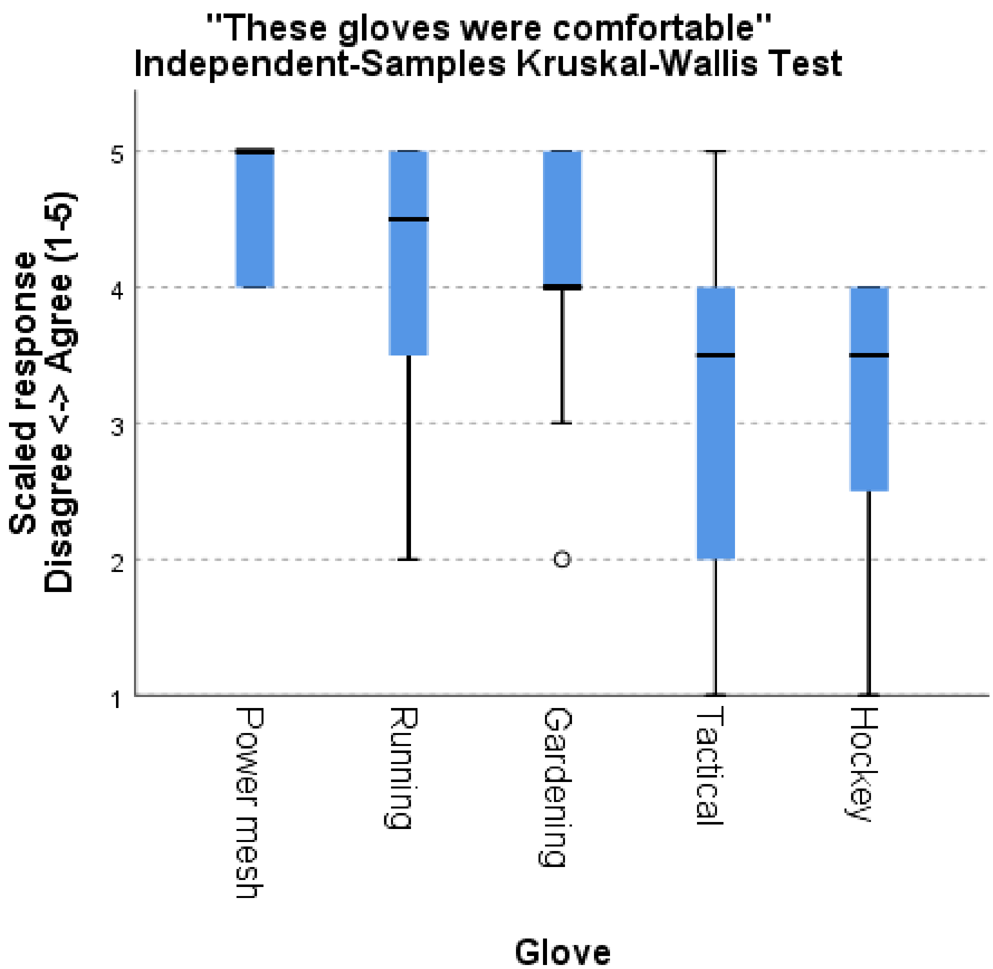


B.4.4 "These gloves were comfortable"

Independent-Samples Kruskal-Wallis Test Summary

Total N	60
Test Statistic	17.365[a]
Degree Of Freedom	4
Asymptotic Sig.(2-sided test)	.002

a. The test statistic is adjusted for ties.



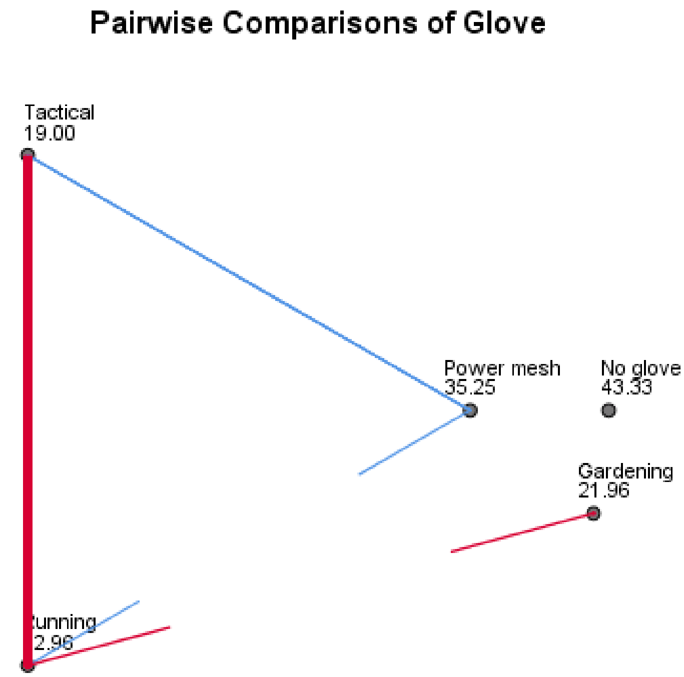
Pairwise Comparisons of Glove

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Hockey-Tactical	2.958	6.775	.437	.662	1.000
Hockey-Gardening	13.958	6.775	2.060	.039	.590
Hockey-Running	16.250	6.775	2.399	.016	.247
Hockey-Power mesh	24.333	6.775	3.592	.000	.005
Tactical-Gardening	11.000	6.775	1.624	.104	1.000
Tactical-Running	13.292	6.775	1.962	.050	.747
Tactical-Power mesh	21.375	6.775	3.155	.002	.024
Gardening-Running	2.292	6.775	.338	.735	1.000
Gardening-Power mesh	10.375	6.775	1.531	.126	1.000
Running-Power mesh	8.083	6.775	1.193	.233	1.000

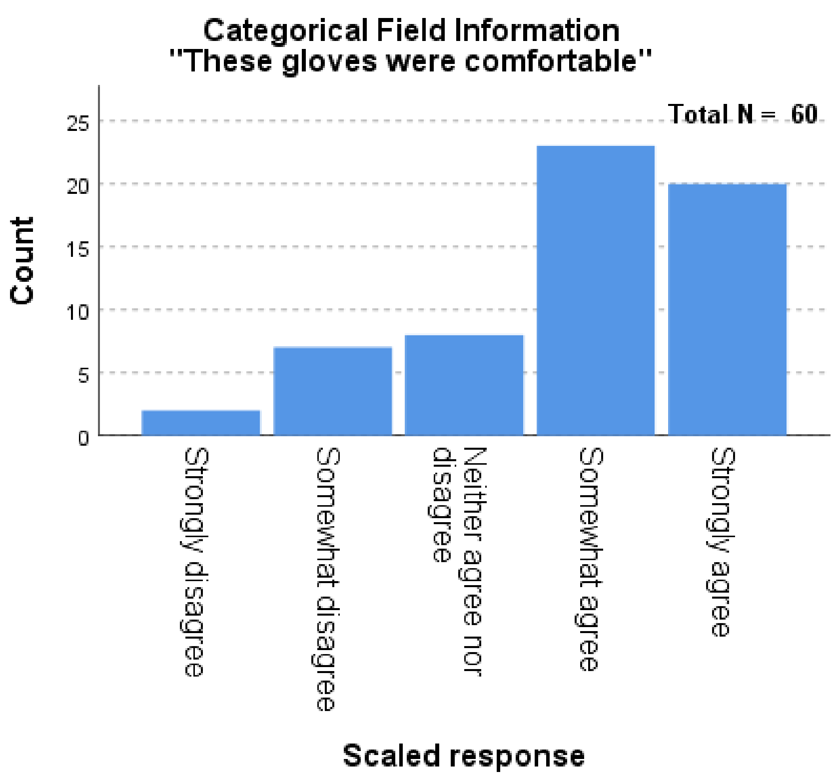
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



Each node shows the sample average rank of Glove.

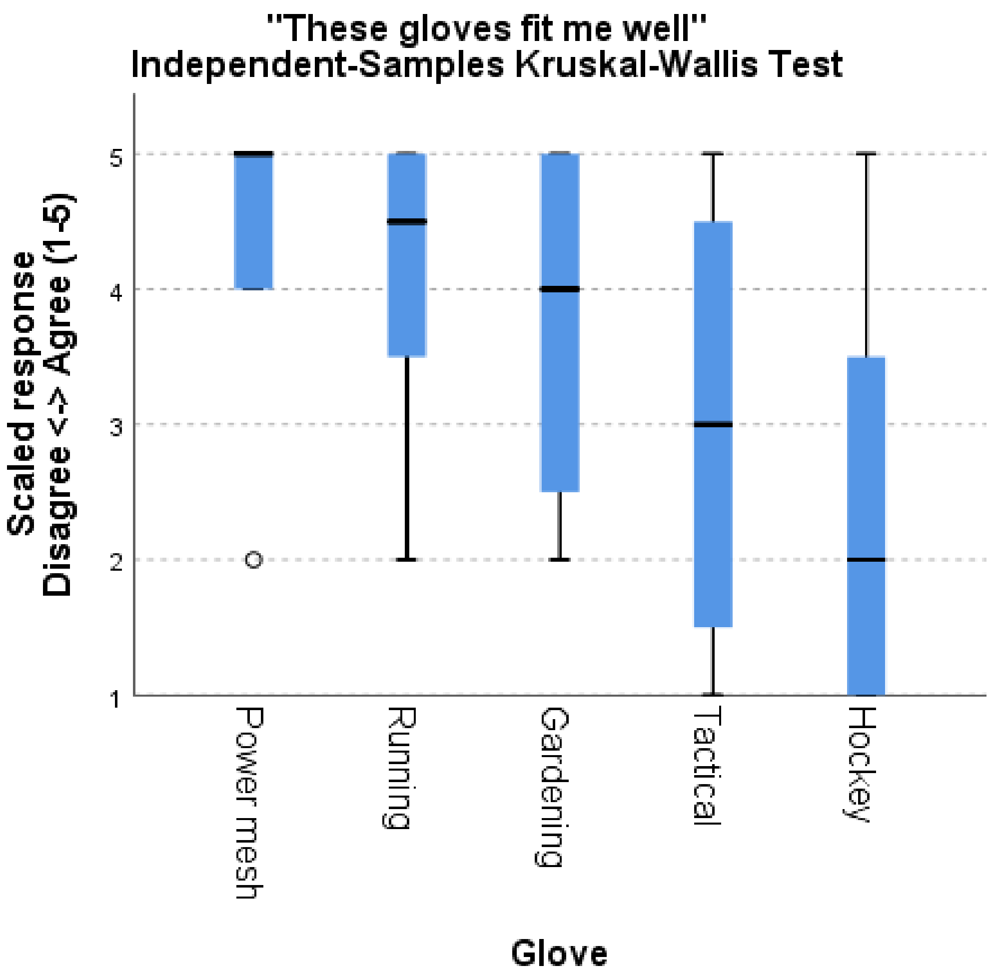


B.4.5 "These gloves fit me well"

Independent-Samples Kruskal-Wallis Test Summary

Total N	60
Test Statistic	16.469[a]
Degree Of Freedom	4
Asymptotic Sig.(2-sided test)	.002

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

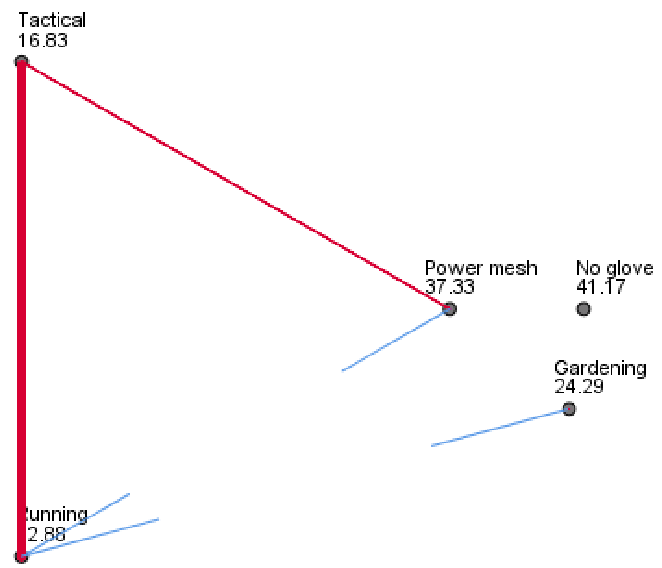
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Hockey-Tactical	7.458	6.895	1.082	.279	1.000
Hockey-Gardening	16.042	6.895	2.327	.020	.300
Hockey-Running	20.500	6.895	2.973	.003	.044
Hockey-Power mesh	24.333	6.895	3.529	.000	.006
Tactical-Gardening	8.583	6.895	1.245	.213	1.000
Tactical-Running	13.042	6.895	1.892	.059	.878
Tactical-Power mesh	16.875	6.895	2.448	.014	.216
Gardening-Running	4.458	6.895	.647	.518	1.000
Gardening-Power mesh	8.292	6.895	1.203	.229	1.000
Running-Power mesh	3.833	6.895	.556	.578	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

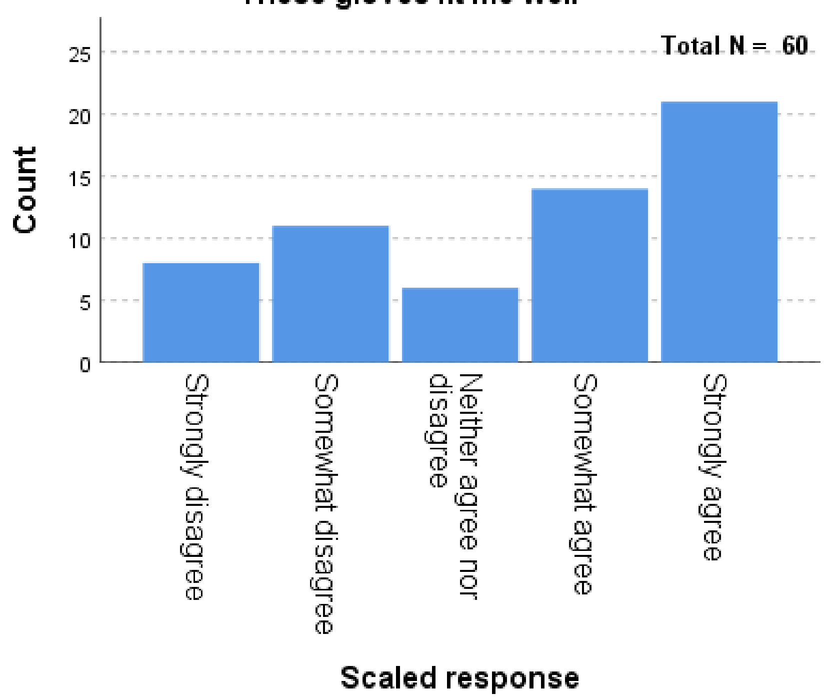
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.

Categorical Field Information
"These gloves fit me well"

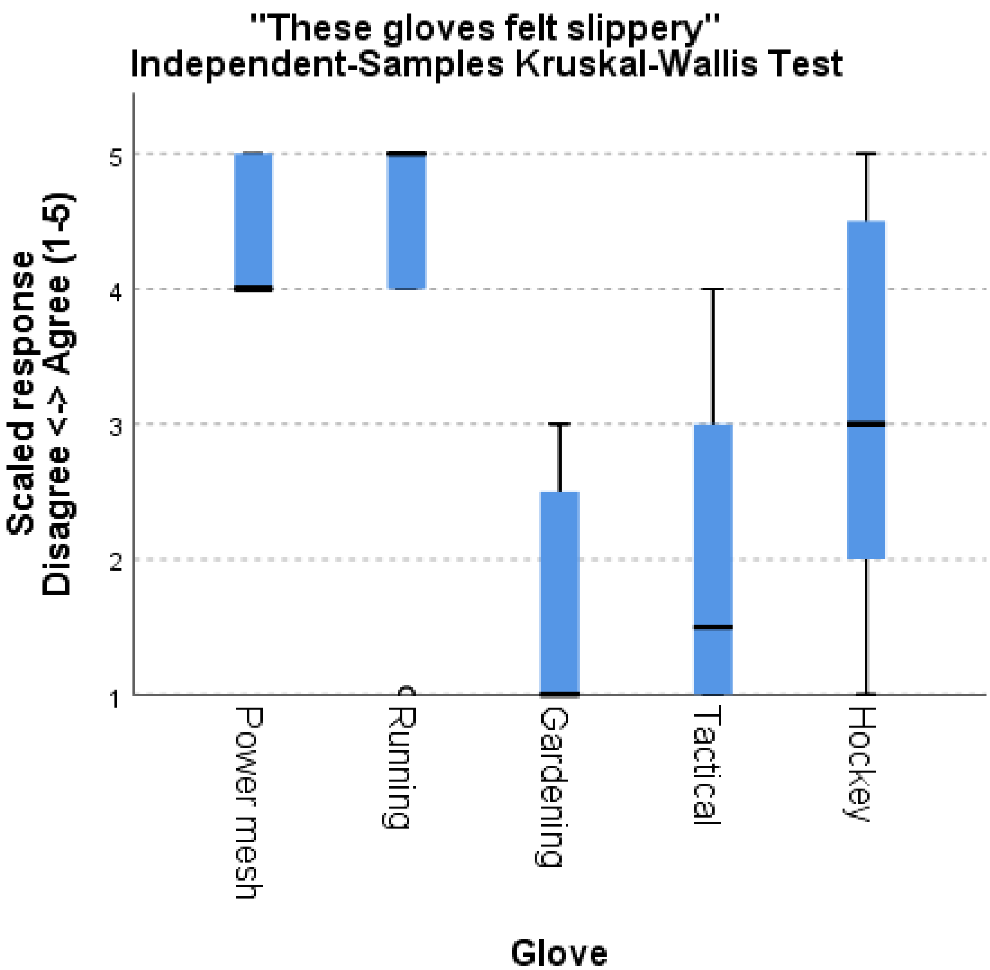


B.4.6 “These gloves felt slippery”

Independent-Samples Kruskal-Wallis Test Summary

Total N	60
Test Statistic	30.376[a]
Degree Of Freedom	4
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



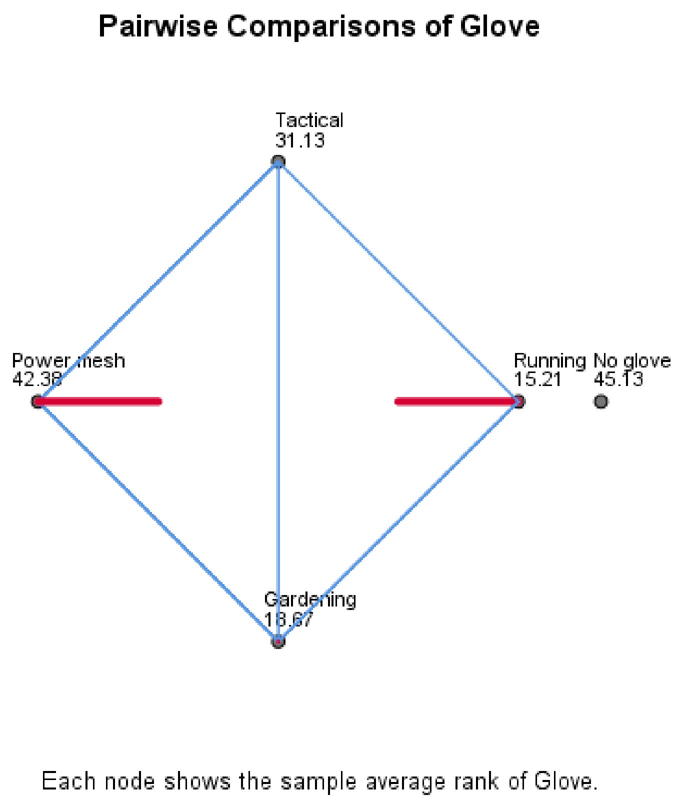
Pairwise Comparisons of Glove

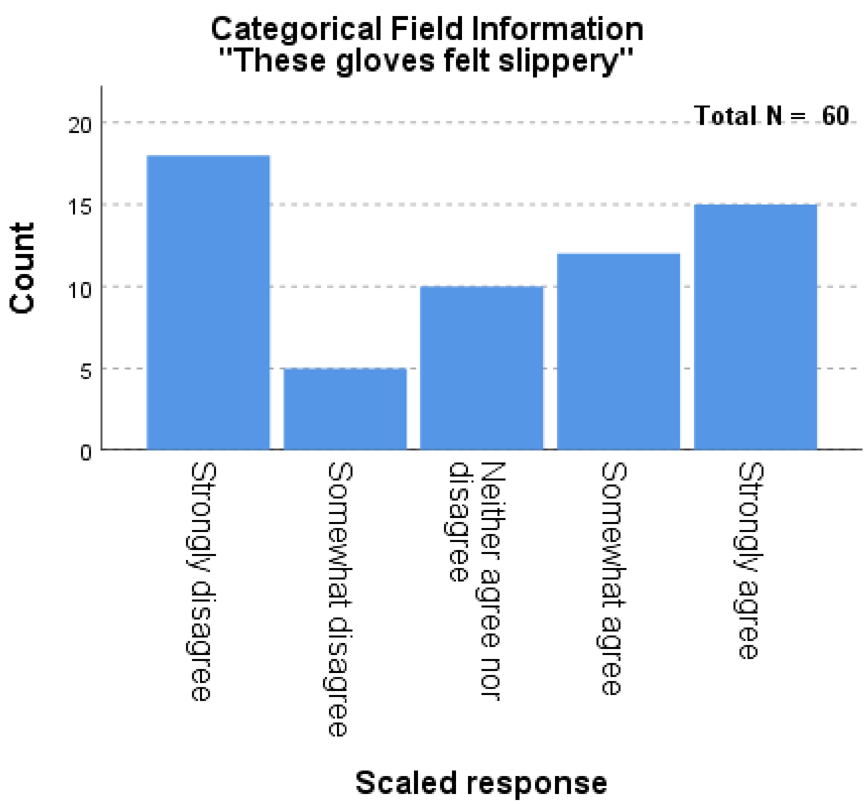
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Gardening-Tactical	-3.458	6.929	-.499	.618	1.000
Gardening-Hockey	-15.917	6.929	-2.297	.022	.324
Gardening-Running	27.167	6.929	3.921	.000	.001
Gardening-Power mesh	29.917	6.929	4.318	.000	.000
Tactical-Hockey	-12.458	6.929	-1.798	.072	1.000
Tactical-Running	23.708	6.929	3.422	.001	.009
Tactical-Power mesh	26.458	6.929	3.819	.000	.002
Hockey-Running	11.250	6.929	1.624	.104	1.000
Hockey-Power mesh	14.000	6.929	2.021	.043	.650
Running-Power mesh	2.750	6.929	.397	.691	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.


Pairwise Comparisons of Glove
Tactical
31.13
Power mesh
42.38
Running
15.21
No glove
45.13
Gardening
18.67
Each node shows the sample average rank of Glove.

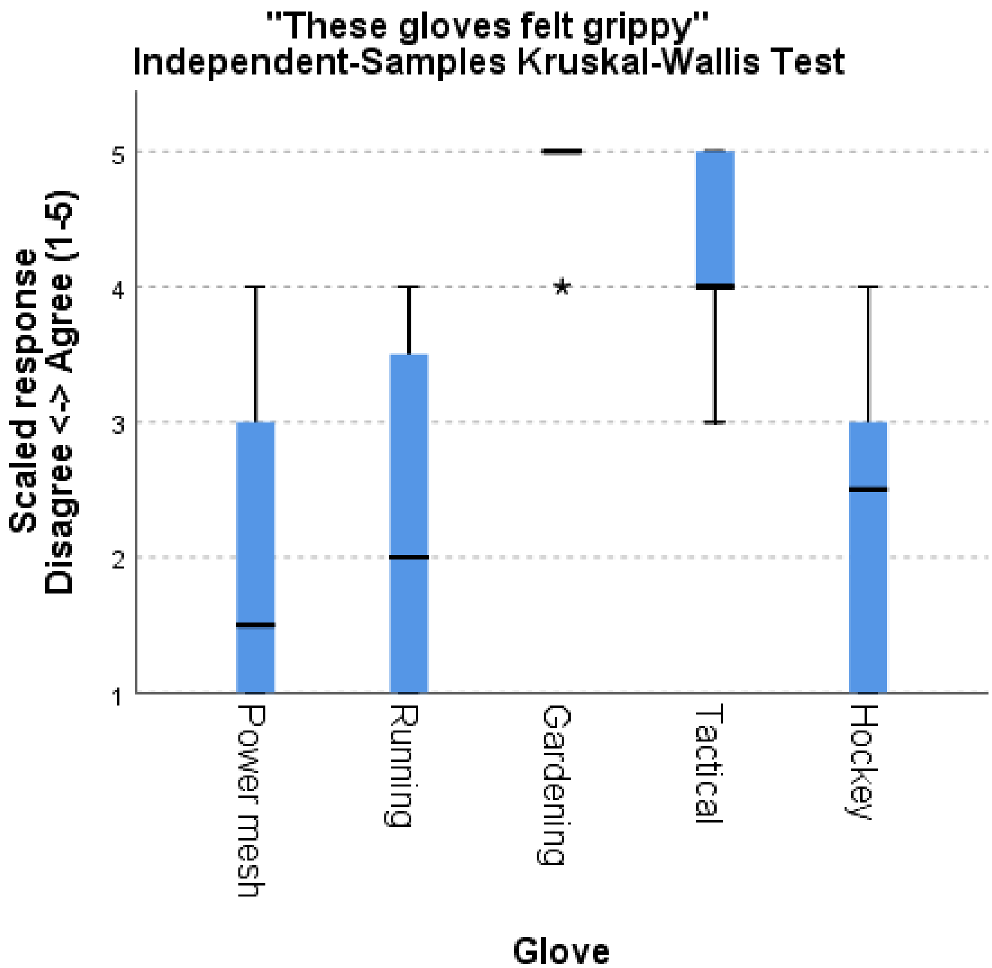

Categorical Field Information
"These gloves felt slippery"
Total N = 60
Count
20
15
10
5
0
Strongly disagree
Somewhat disagree
Neither agree nor disagree
Somewhat agree
Strongly agree
Scaled response

B.4.7 "These gloves felt grippy"

Independent-Samples Kruskal-Wallis Test Summary

Total N	60
Test Statistic	39.816[a]
Degree Of Freedom	4
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



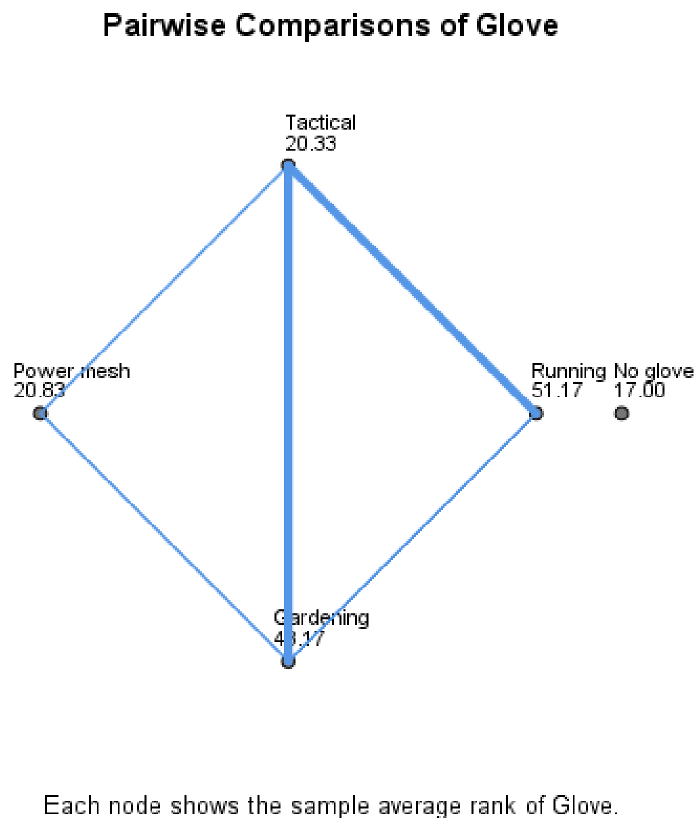
Pairwise Comparisons of Glove

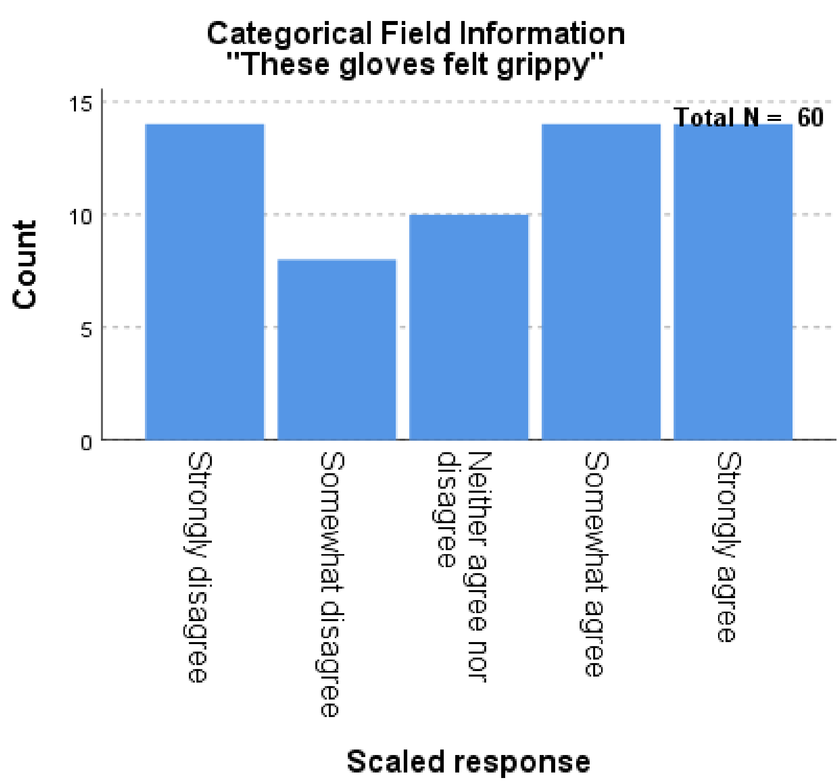
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Power mesh-Hockey	-3.333	6.968	-.478	.632	1.000
Power mesh-Running	-3.833	6.968	-.550	.582	1.000
Power mesh-Tactical	-26.167	6.968	-3.755	.000	.003
Power mesh-Gardening	-34.167	6.968	-4.903	.000	.000
Hockey-Running	.500	6.968	.072	.943	1.000
Hockey-Tactical	22.833	6.968	3.277	.001	.016
Hockey-Gardening	30.833	6.968	4.425	.000	.000
Running-Tactical	-22.333	6.968	-3.205	.001	.020
Running-Gardening	-30.333	6.968	-4.353	.000	.000
Tactical-Gardening	8.000	6.968	1.148	.251	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.


Pairwise Comparisons of Glove
Tactical
20.33
Power mesh
20.83
Running
51.17
No glove
17.00
Gardening
43.17
Each node shows the sample average rank of Glove.

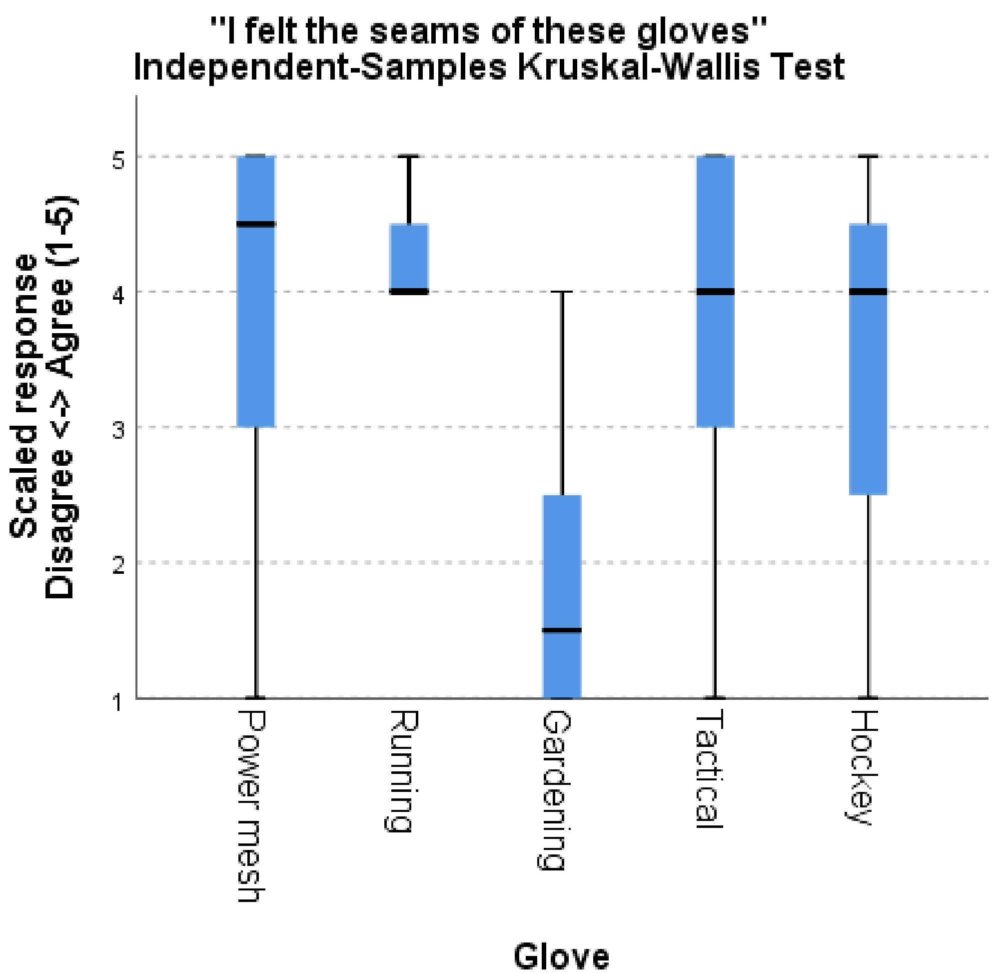

Categorical Field Information
"These gloves felt grippy"
Total N = 60
Count
15
10
5
0
Strongly disagree
Somewhat disagree
Neither agree nor disagree
Somewhat agree
Strongly agree
Scaled response

B.4.8 "I felt the seams of these gloves"

Independent-Samples Kruskal-Wallis Test Summary

Total N	60
Test Statistic	18.701[a]
Degree Of Freedom	4
Asymptotic Sig.(2-sided test)	.001

a. The test statistic is adjusted for ties.



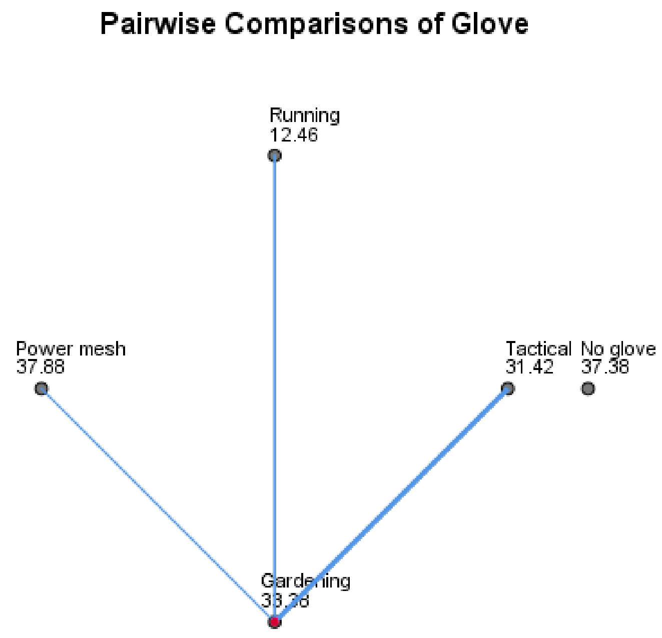
Pairwise Comparisons of Glove

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Gardening-Hockey	-18.958	6.831	-2.776	.006	.083
Gardening-Tactical	-20.917	6.831	-3.062	.002	.033
Gardening-Power mesh	24.917	6.831	3.648	.000	.004
Gardening-Running	25.417	6.831	3.721	.000	.003
Hockey-Tactical	1.958	6.831	.287	.774	1.000
Hockey-Power mesh	5.958	6.831	.872	.383	1.000
Hockey-Running	6.458	6.831	.946	.344	1.000
Tactical-Power mesh	4.000	6.831	.586	.558	1.000
Tactical-Running	4.500	6.831	.659	.510	1.000
Power mesh-Running	-.500	6.831	-.073	.942	1.000

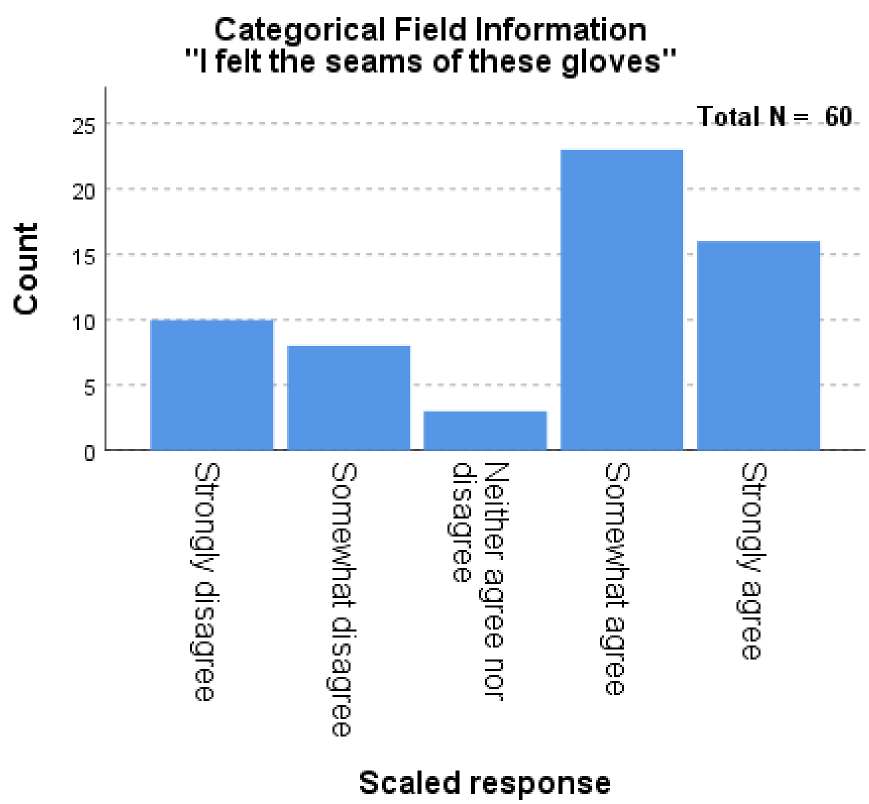
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



Each node shows the sample average rank of Glove.

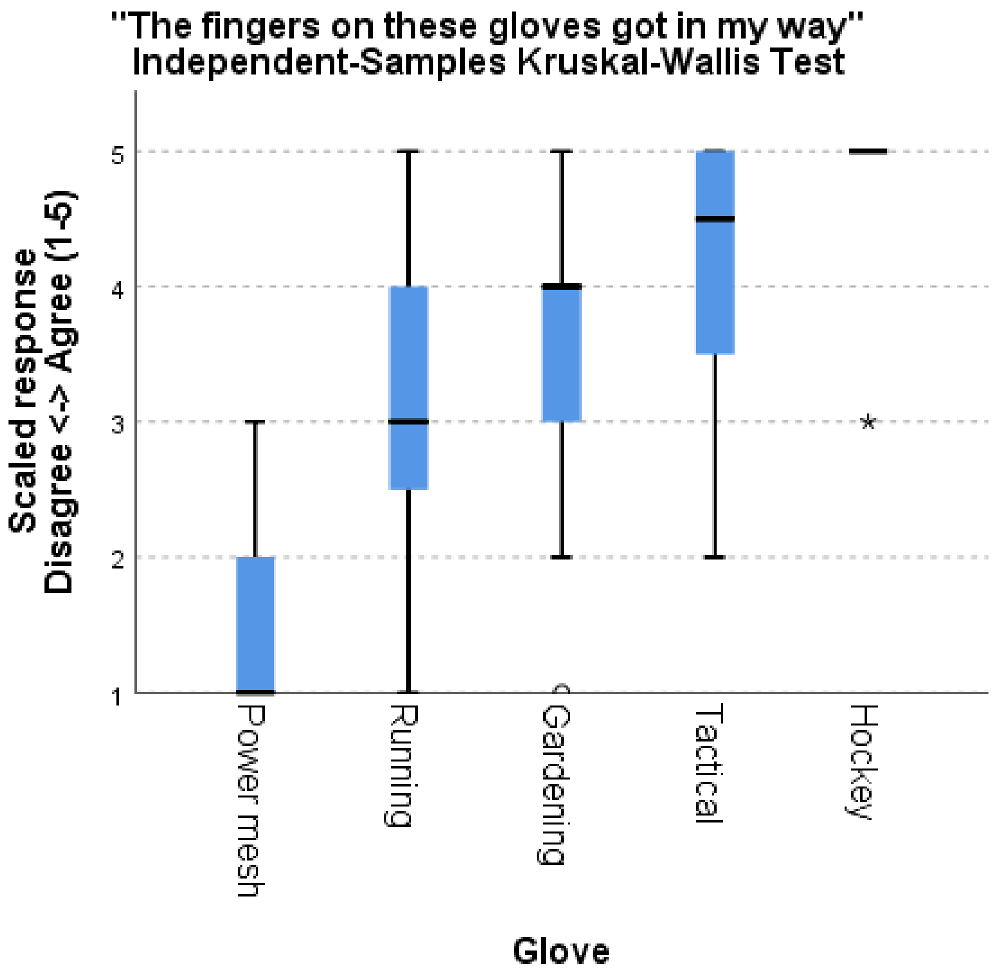


B.4.9 "The fingers on these gloves got in my way"

Independent-Samples Kruskal-Wallis Test Summary

Total N	60
Test Statistic	33.738[a]
Degree Of Freedom	4
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

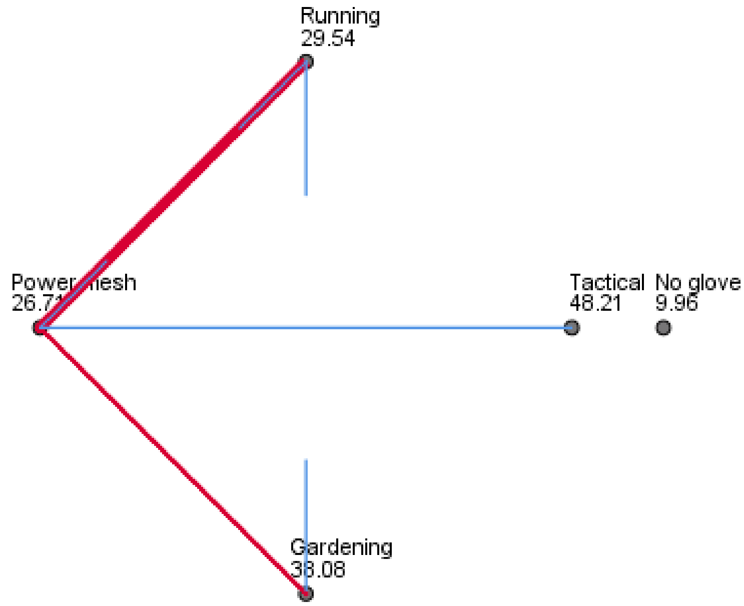
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Power mesh-Running	-16.750	6.922	-2.420	.016	.233
Power mesh-Gardening	-19.583	6.922	-2.829	.005	.070
Power mesh-Tactical	-28.125	6.922	-4.063	.000	.001
Power mesh-Hockey	-38.250	6.922	-5.526	.000	.000
Running-Gardening	-2.833	6.922	-.409	.682	1.000
Running-Tactical	-11.375	6.922	-1.643	.100	1.000
Running-Hockey	-21.500	6.922	-3.106	.002	.028
Gardening-Tactical	-8.542	6.922	-1.234	.217	1.000
Gardening-Hockey	-18.667	6.922	-2.697	.007	.105
Tactical-Hockey	-10.125	6.922	-1.463	.144	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

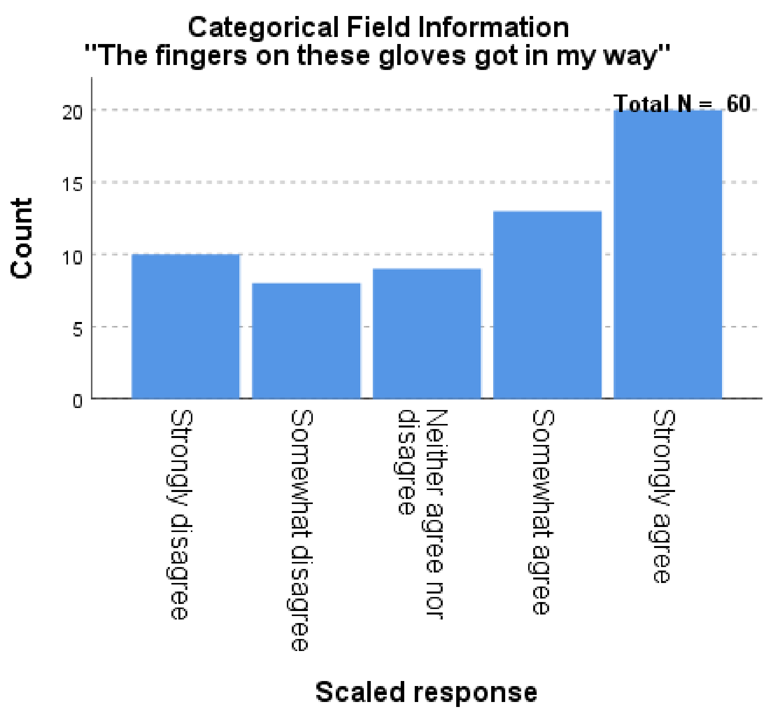
Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



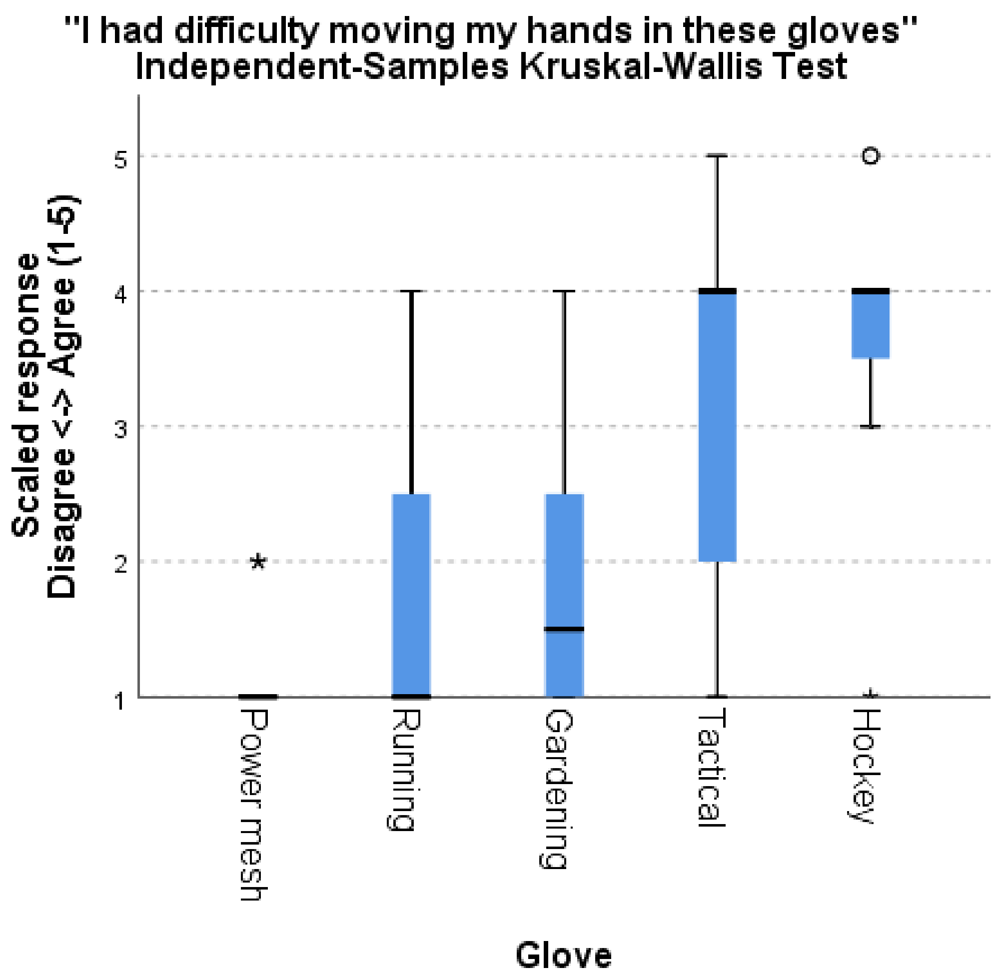
Each node shows the sample average rank of Glove.



B.4.10 "I had difficulty moving my hands in these gloves"

Independent-Samples Kruskal-Wallis Test Summary	
Total N	60
Test Statistic	30.048[a]
Degree Of Freedom	4
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



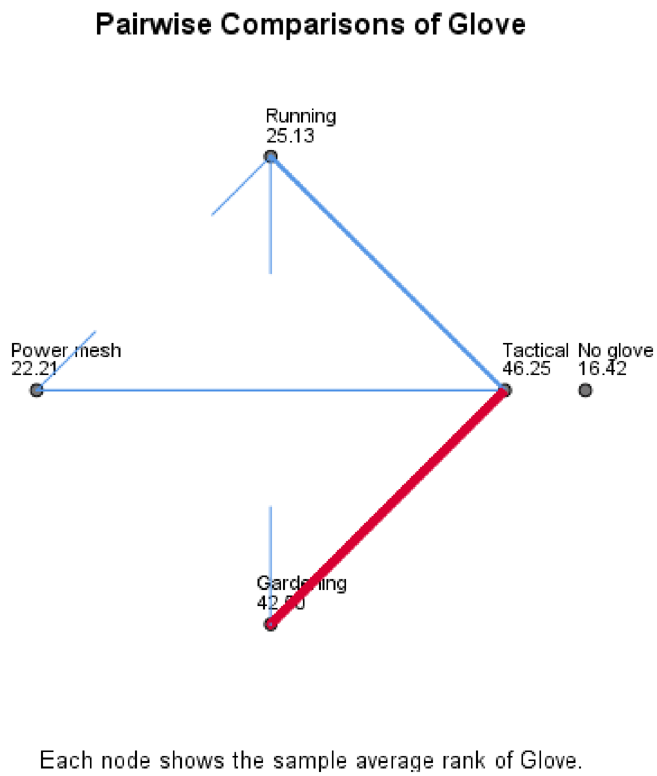
Pairwise Comparisons of Glove

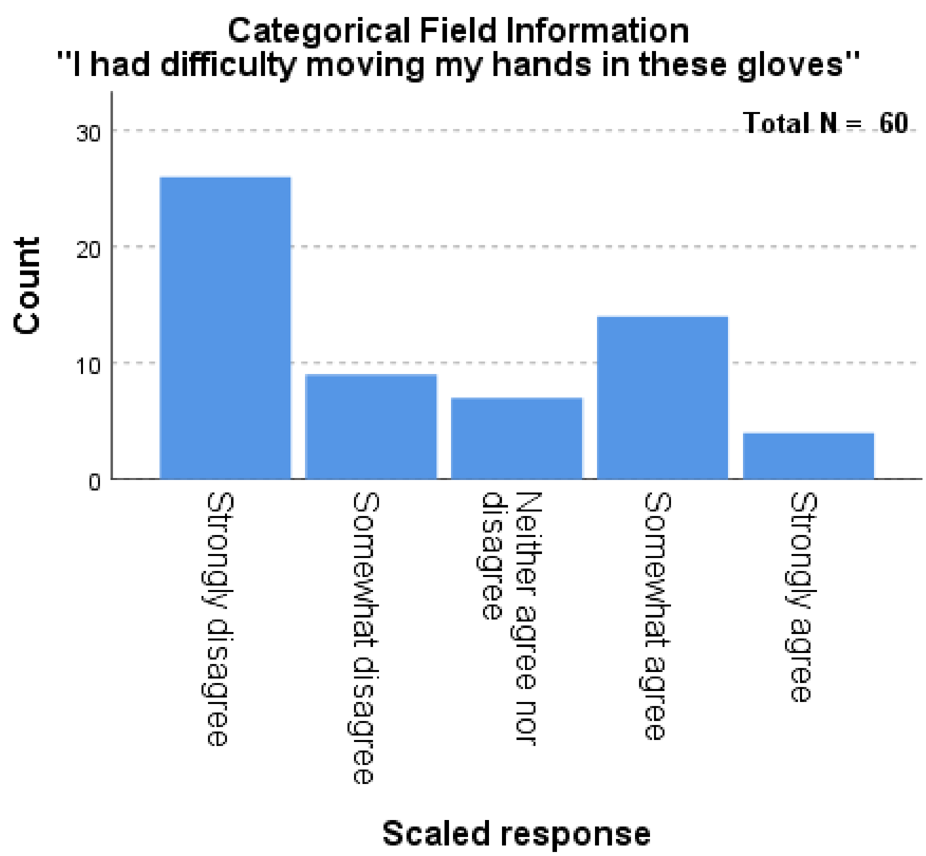
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Power mesh-Running	-5.792	6.767	-.856	.392	1.000
Power mesh-Gardening	-8.708	6.767	-1.287	.198	1.000
Power mesh-Tactical	-26.083	6.767	-3.854	.000	.002
Power mesh-Hockey	-29.833	6.767	-4.408	.000	.000
Running-Gardening	-2.917	6.767	-.431	.666	1.000
Running-Tactical	-20.292	6.767	-2.998	.003	.041
Running-Hockey	-24.042	6.767	-3.553	.000	.006
Gardening-Tactical	-17.375	6.767	-2.567	.010	.154
Gardening-Hockey	-21.125	6.767	-3.122	.002	.027
Tactical-Hockey	-3.750	6.767	-.554	.579	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.


Pairwise Comparisons of Glove
Running
25.13
Power mesh
22.21
Tactical
46.25
No glove
16.42
Gardening
42.50
Each node shows the sample average rank of Glove.

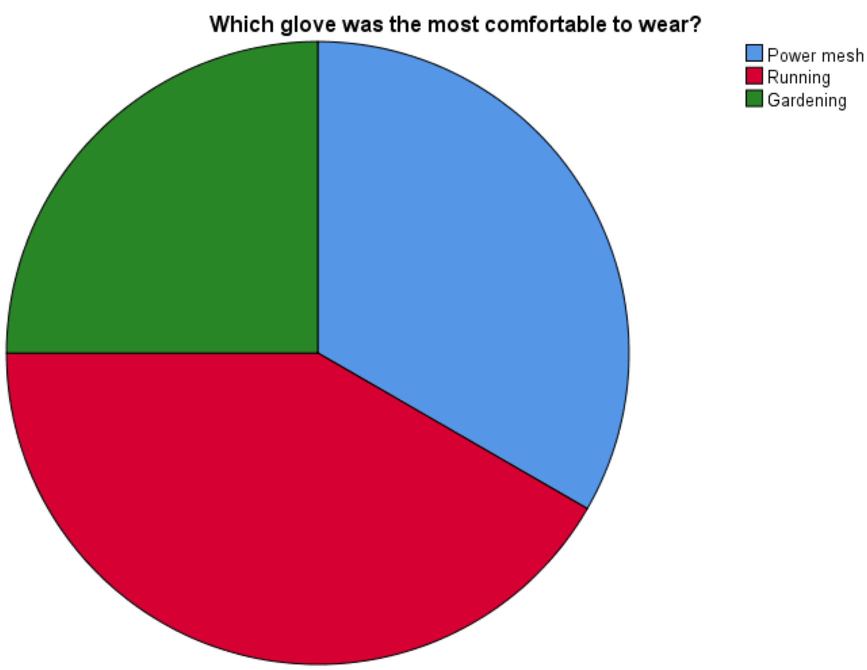

Categorical Field Information
"I had difficulty moving my hands in these gloves"
Total N = 60
Count
30
20
10
0
Strongly disagree
Somewhat disagree
Neither agree nor disagree
Somewhat agree
Strongly agree
Scaled response

B.5 Glove preference data

B.5.1 Which glove was the most comfortable to wear?

Which glove was the most comfortable to wear?

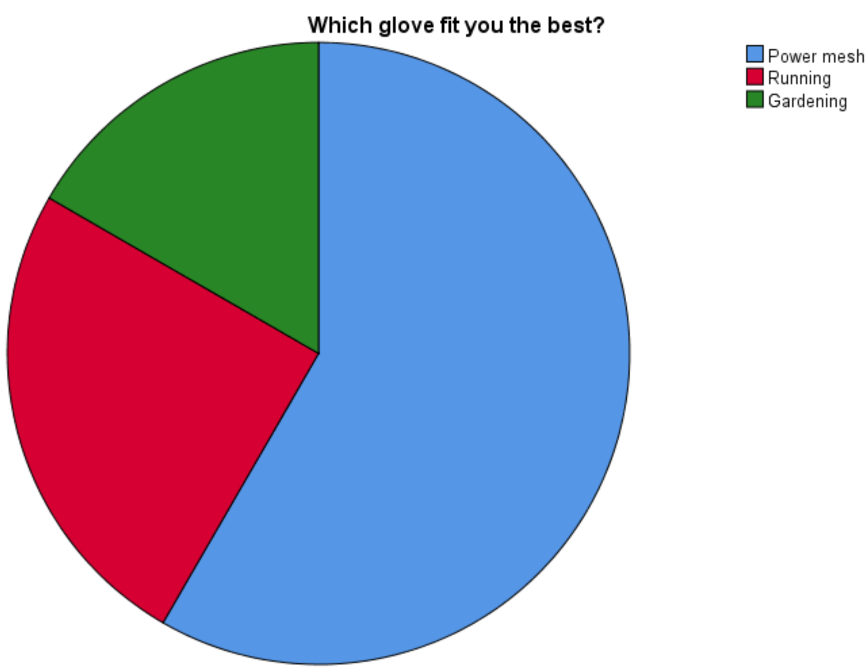
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Power mesh	4	33.3	33.3	33.3
	Running	5	41.7	41.7	75.0
	Gardening	3	25.0	25.0	100.0
	Total	12	100.0	100.0	



B.5.2 Which glove fit you the best?

Which glove fit you the best?

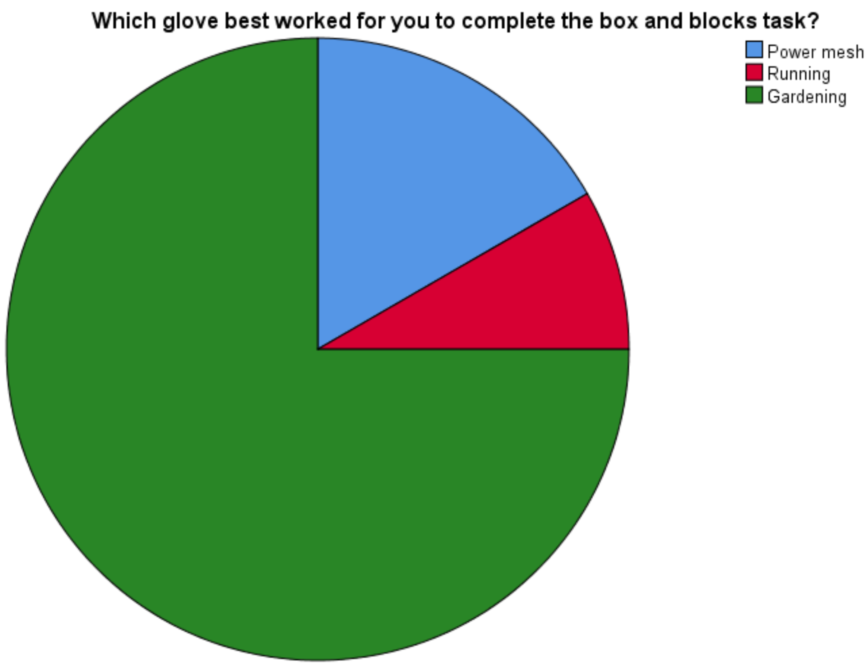
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Power mesh	7	58.3	58.3	58.3
	Running	3	25.0	25.0	83.3
	Gardening	2	16.7	16.7	100.0
	Total	12	100.0	100.0	



B.5.3 Which glove best worked for you to complete the box and blocks task?

Which glove best worked for you to complete the box and blocks task?

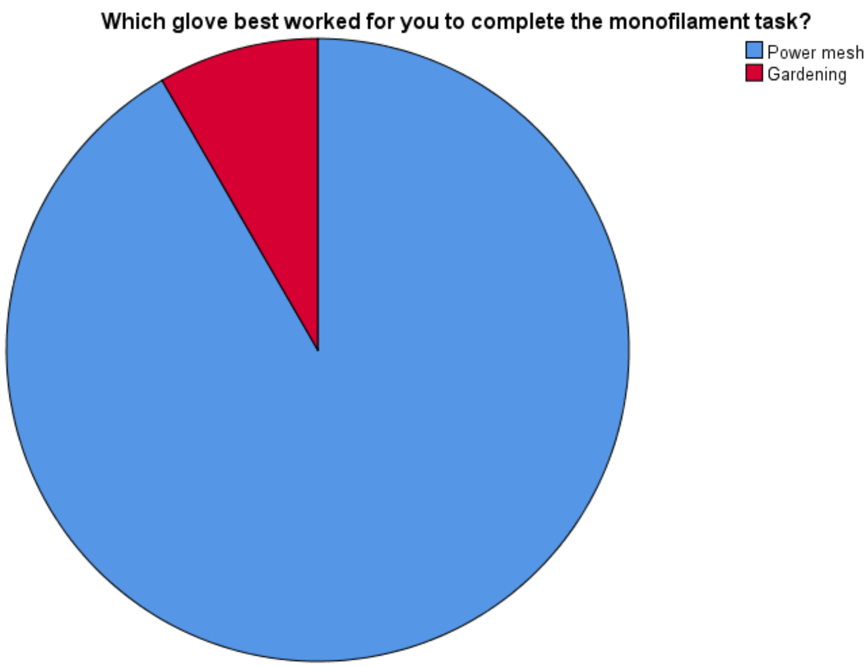
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Power mesh	2	16.7	16.7	16.7
	Running	1	8.3	8.3	25.0
	Gardening	9	75.0	75.0	100.0
	Total	12	100.0	100.0	



B.5.4 Which glove best worked for you to complete the monofilament task?

Which glove best worked for you to complete the monofilament task?

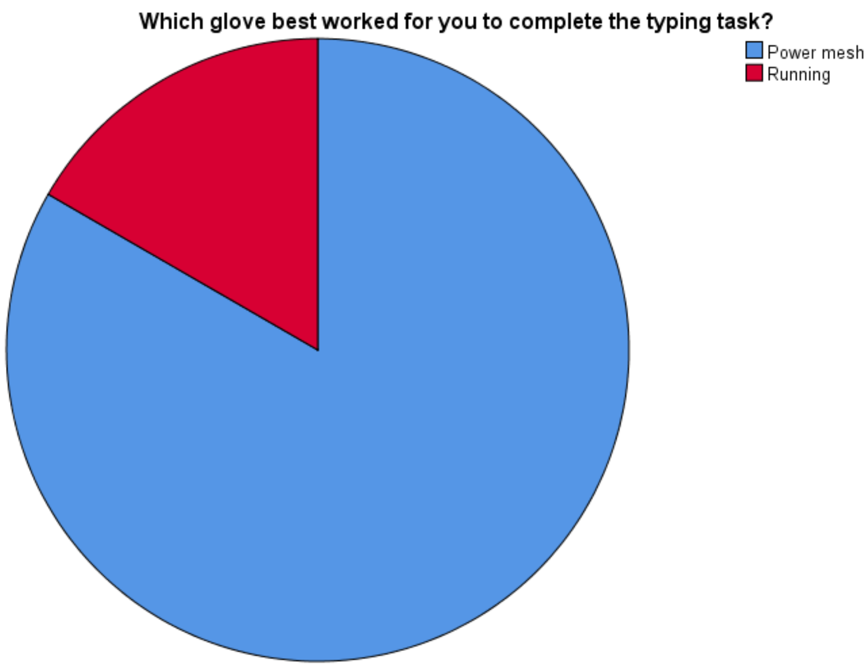
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Power mesh	11	91.7	91.7	91.7
	Gardening	1	8.3	8.3	100.0
	Total	12	100.0	100.0	



B.5.5 Which glove best worked for you to complete the typing task?

Which glove best worked for you to complete the typing task?

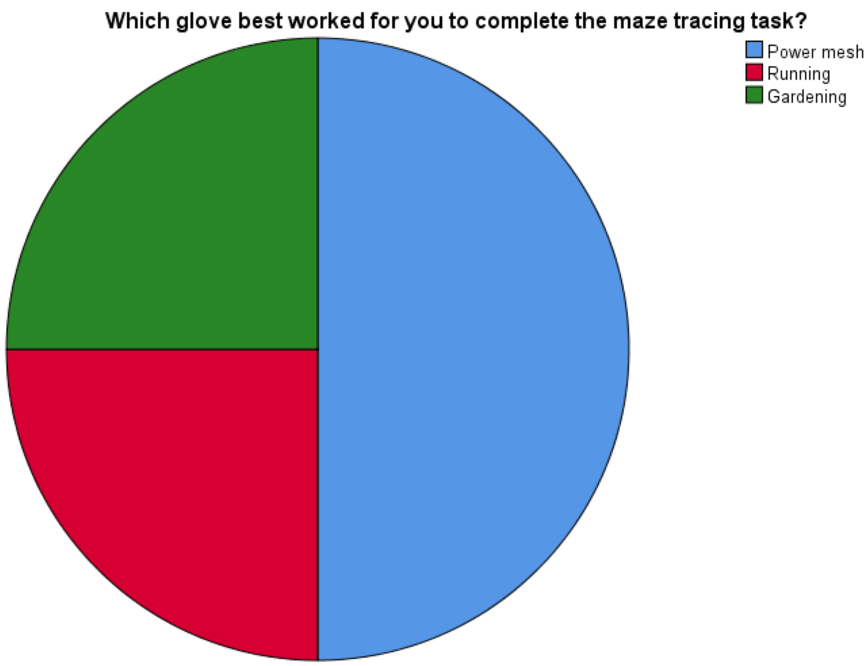
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Power mesh	10	83.3	83.3	83.3
	Running	2	16.7	16.7	100.0
	Total	12	100.0	100.0	



B.5.6 Which glove best worked for you to complete the maze tracing task?

Which glove best worked for you to complete the maze tracing task?

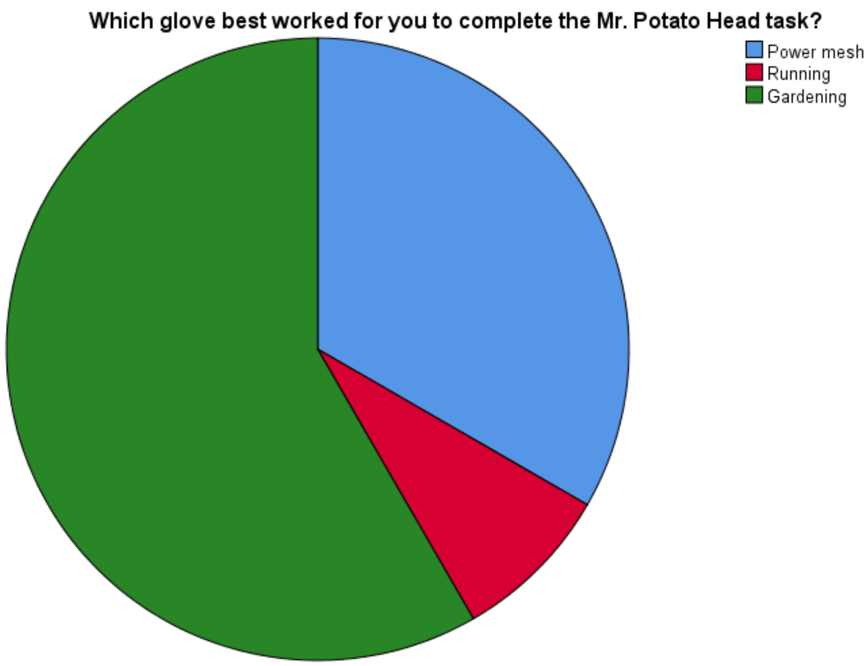
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Power mesh	6	50.0	50.0	50.0
	Running	3	25.0	25.0	75.0
	Gardening	3	25.0	25.0	100.0
	Total	12	100.0	100.0	



B.5.7 Which glove best worked for you to complete the Mr. Potato Head task?

Which glove best worked for you to complete the Mr. Potato Head task?

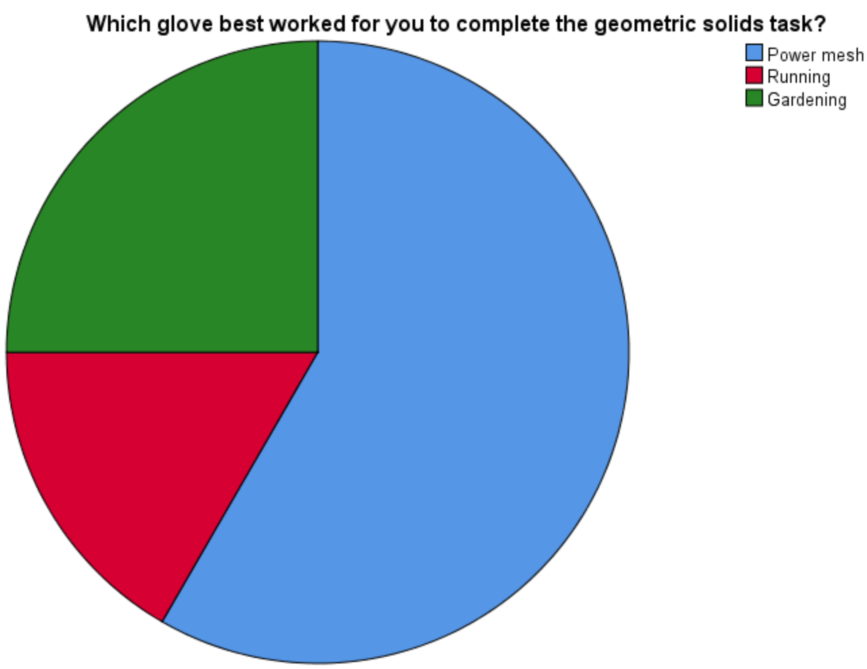
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Power mesh	4	33.3	33.3	33.3
	Running	1	8.3	8.3	41.7
	Gardening	7	58.3	58.3	100.0
	Total	12	100.0	100.0	



B.5.8 Which glove best worked for you to complete the geometric solids task?

Which glove best worked for you to complete the geometric solids task?

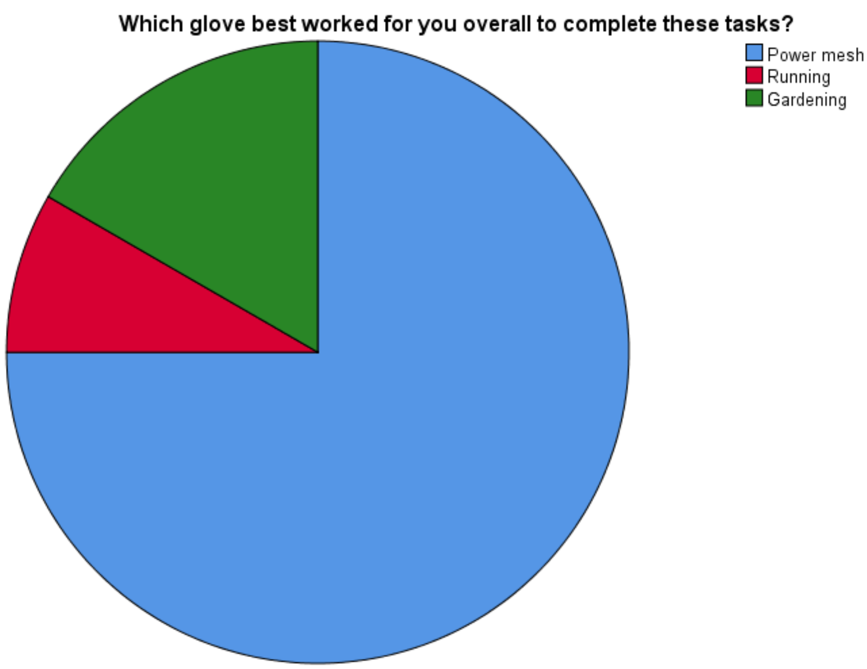
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Power mesh	7	58.3	58.3	58.3
	Running	2	16.7	16.7	75.0
	Gardening	3	25.0	25.0	100.0
	Total	12	100.0	100.0	



B.5.9 Which glove best worked for you overall to complete these tasks?

Which glove best worked for you overall to complete these tasks?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Power mesh	9	75.0	75.0	75.0
	Running	1	8.3	8.3	83.3
	Gardening	2	16.7	16.7	100.0
	Total	12	100.0	100.0	



APPENDIX C. STUDY 2: COMMERCIAL GLOVE DATA

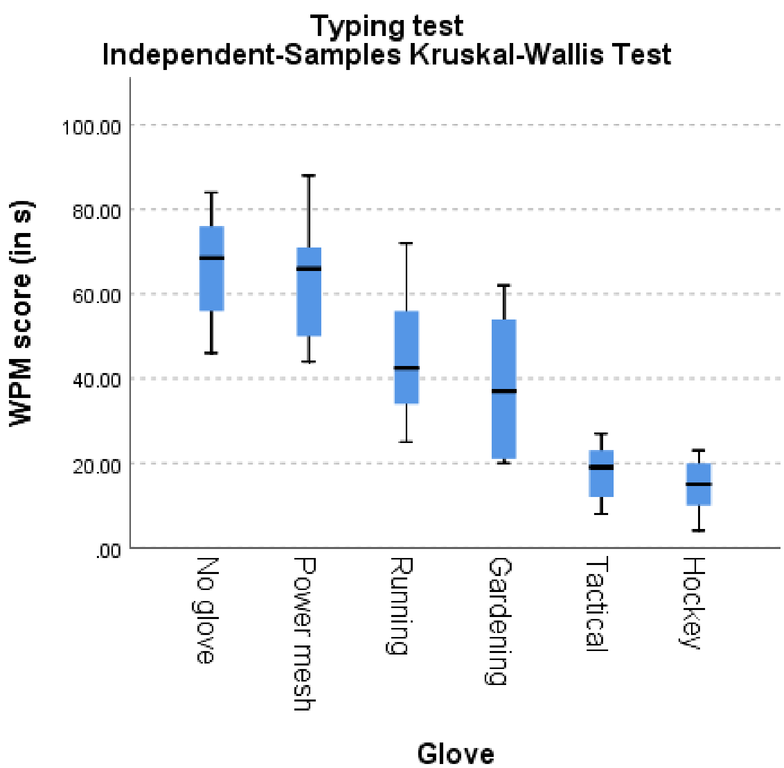
C.1 Task performance data

C.1.1 Typing test

Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	26.048[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



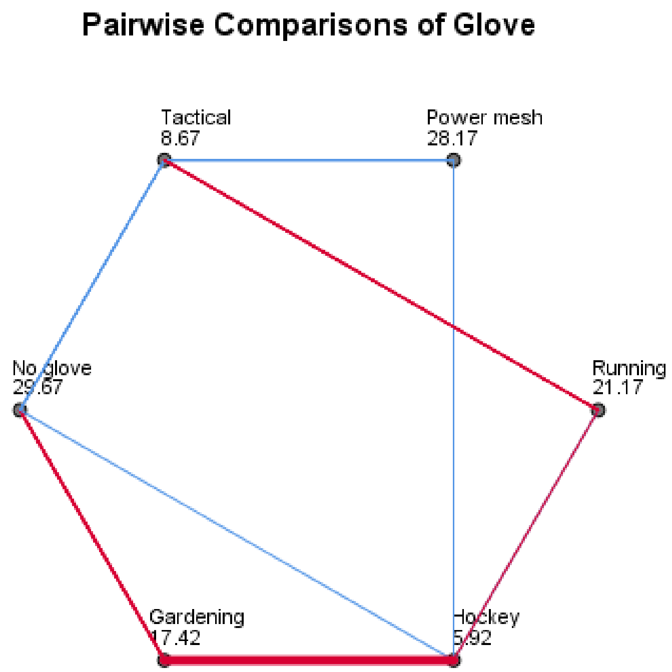
Pairwise Comparisons of Glove

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Hockey-Tactical	2.750	6.080	.452	.651	1.000
Hockey-Gardening	11.500	6.080	1.891	.059	.878
Hockey-Running	15.250	6.080	2.508	.012	.182
Hockey-Power mesh	22.250	6.080	3.660	.000	.004
Hockey-No glove	23.750	6.080	3.906	.000	.001
Tactical-Gardening	8.750	6.080	1.439	.150	1.000
Tactical-Running	12.500	6.080	2.056	.040	.597
Tactical-Power mesh	19.500	6.080	3.207	.001	.020
Tactical-No glove	21.000	6.080	3.454	.001	.008
Gardening-Running	3.750	6.080	.617	.537	1.000
Gardening-Power mesh	10.750	6.080	1.768	.077	1.000
Gardening-No glove	12.250	6.080	2.015	.044	.659
Running-Power mesh	7.000	6.080	1.151	.250	1.000
Running-No glove	8.500	6.080	1.398	.162	1.000
Power mesh-No glove	1.500	6.080	.247	.805	1.000

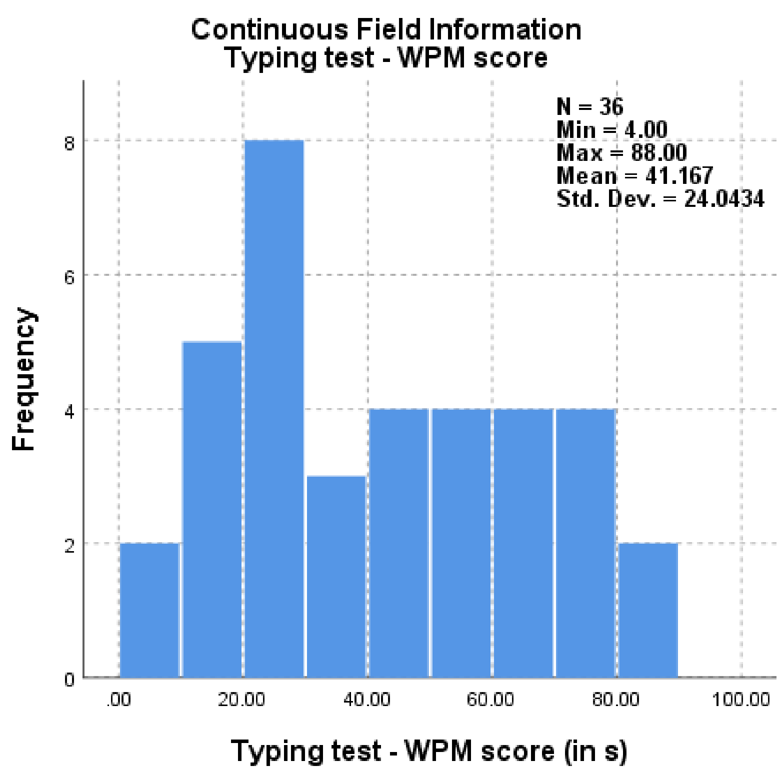
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



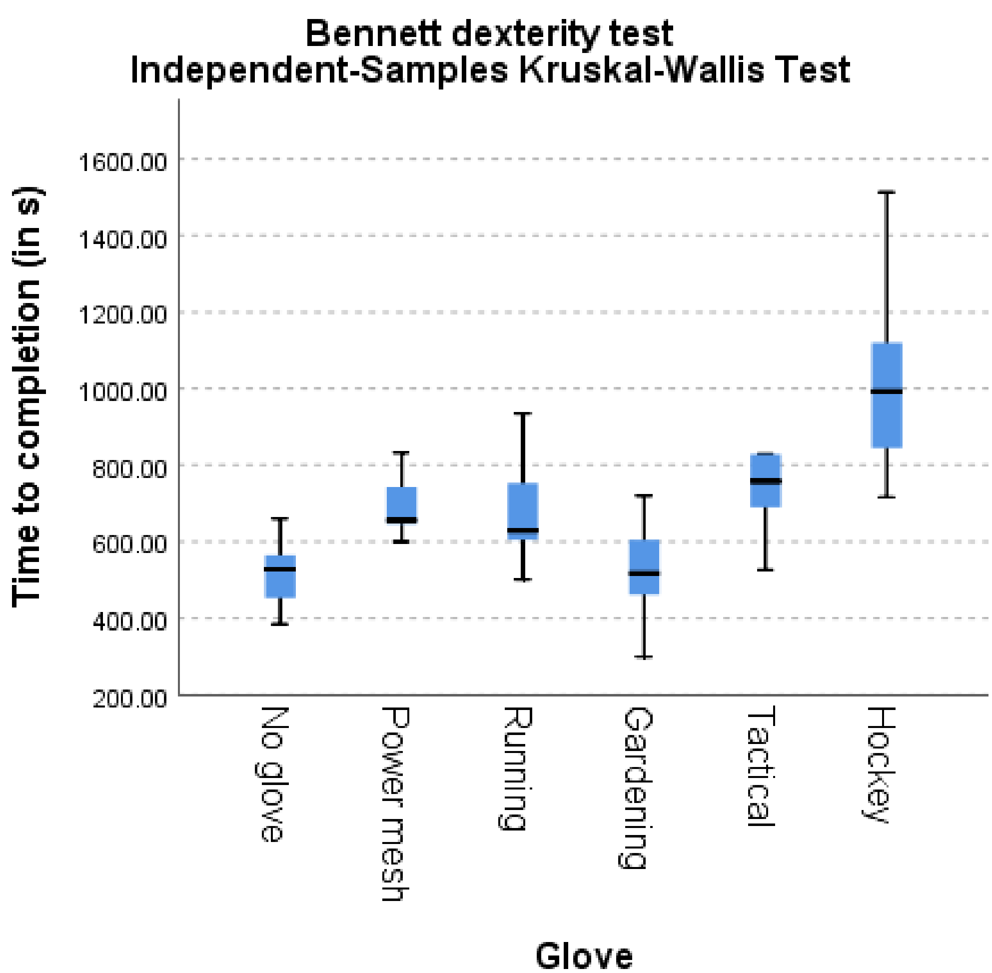
Each node shows the sample average rank of Glove.



C.1.2 Bennett dexterity test

Independent-Samples Kruskal-Wallis Test Summary	
Total N	36
Test Statistic	20.200[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.001

a. The test statistic is adjusted for ties.



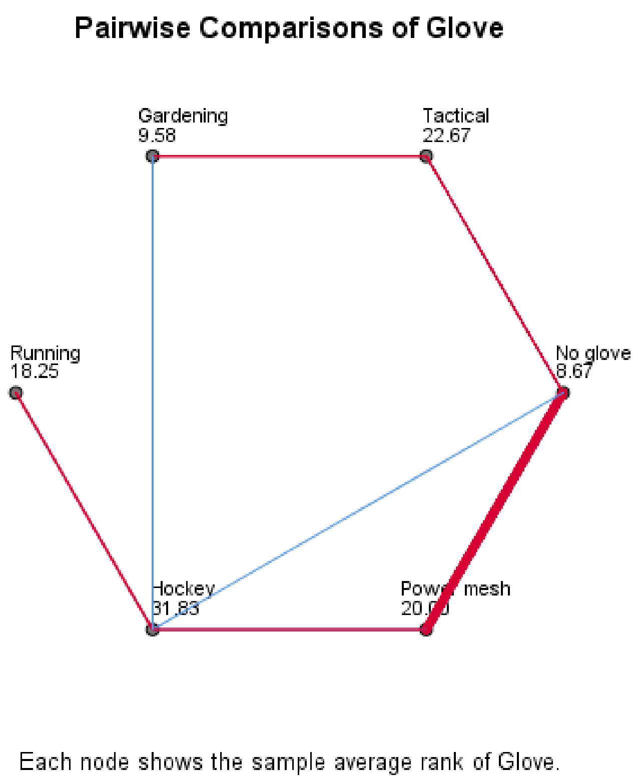
Pairwise Comparisons of Glove

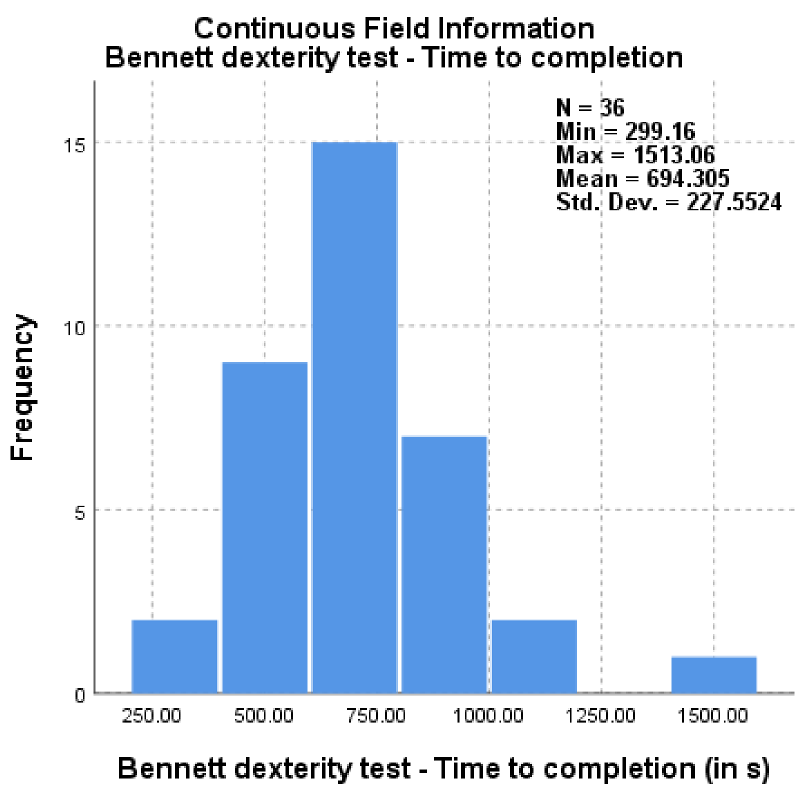
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
No glove-Gardening	-.917	6.082	-.151	.880	1.000
No glove-Running	-9.583	6.082	-1.576	.115	1.000
No glove-Power mesh	-11.333	6.082	-1.863	.062	.936
No glove-Tactical	-14.000	6.082	-2.302	.021	.320
No glove-Hockey	-23.167	6.082	-3.809	.000	.002
Gardening-Running	8.667	6.082	1.425	.154	1.000
Gardening-Power mesh	10.417	6.082	1.713	.087	1.000
Gardening-Tactical	-13.083	6.082	-2.151	.031	.472
Gardening-Hockey	-22.250	6.082	-3.658	.000	.004
Running-Power mesh	1.750	6.082	.288	.774	1.000
Running-Tactical	-4.417	6.082	-.726	.468	1.000
Running-Hockey	-13.583	6.082	-2.233	.026	.383
Power mesh-Tactical	-2.667	6.082	-.438	.661	1.000
Power mesh-Hockey	-11.833	6.082	-1.946	.052	.776
Tactical-Hockey	-9.167	6.082	-1.507	.132	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.


Pairwise Comparisons of Glove
Gardening
9.58
Tactical
22.67
Running
18.25
No glove
8.67
Hockey
31.83
Power mesh
20.00
Each node shows the sample average rank of Glove.

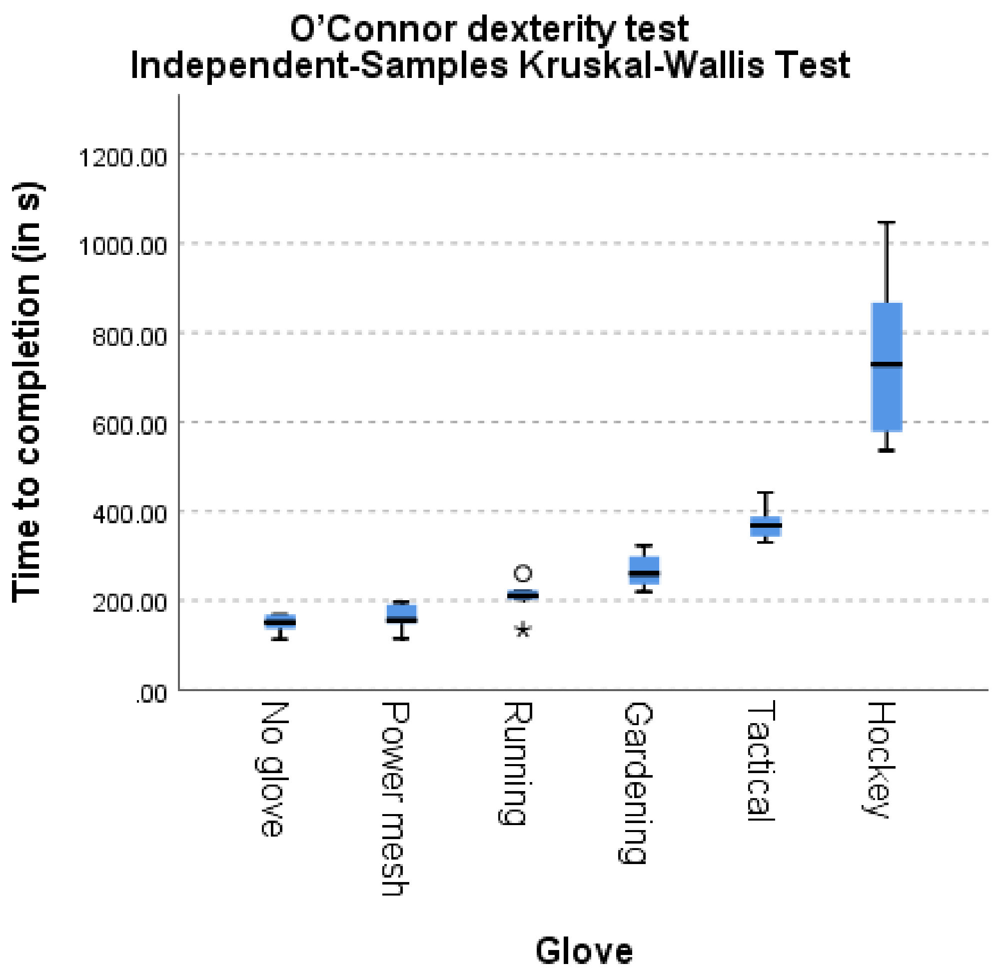

Continuous Field Information
Bennett dexterity test - Time to completion
N = 36
Min = 299.16
Max = 1513.06
Mean = 694.305
Std. Dev. = 227.5524
Frequency
0
5
10
15
250.00
500.00
750.00
1000.00
1250.00
1500.00
Bennett dexterity test - Time to completion (in s)

C.1.3 O'Connor dexterity test

Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	31.255[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



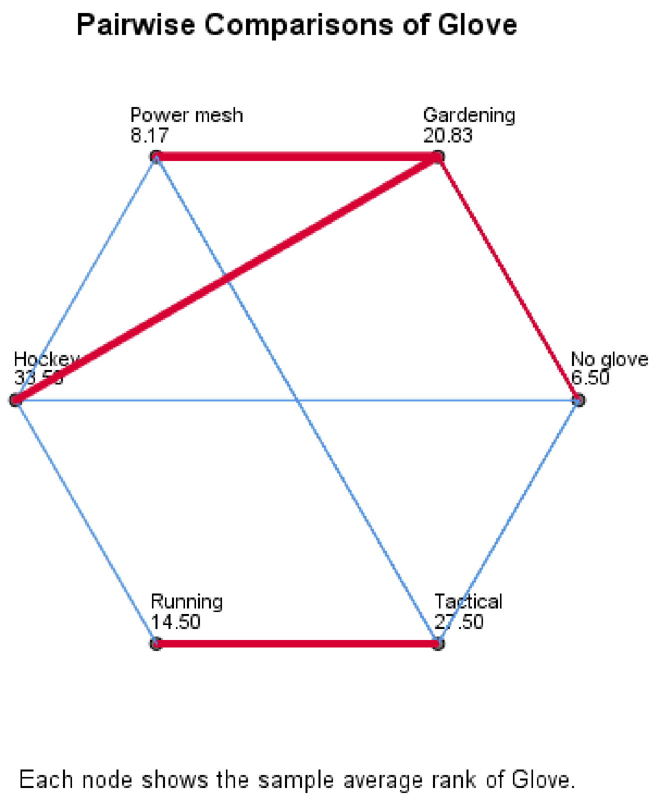
Pairwise Comparisons of Glove

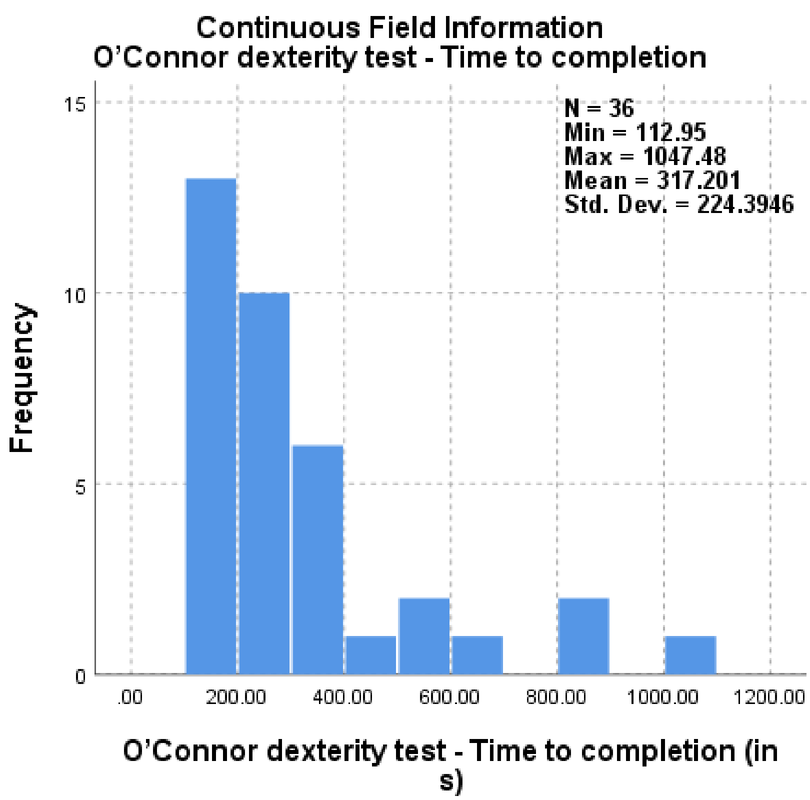
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
No glove-Power mesh	-1.667	6.083	-.274	.784	1.000
No glove-Running	-8.000	6.083	-1.315	.188	1.000
No glove-Gardening	-14.333	6.083	-2.356	.018	.277
No glove-Tactical	-21.000	6.083	-3.452	.001	.008
No glove-Hockey	-27.000	6.083	-4.439	.000	.000
Power mesh-Running	-6.333	6.083	-1.041	.298	1.000
Power mesh-Gardening	-12.667	6.083	-2.082	.037	.560
Power mesh-Tactical	-19.333	6.083	-3.178	.001	.022
Power mesh-Hockey	-25.333	6.083	-4.165	.000	.000
Running-Gardening	-6.333	6.083	-1.041	.298	1.000
Running-Tactical	-13.000	6.083	-2.137	.033	.489
Running-Hockey	-19.000	6.083	-3.124	.002	.027
Gardening-Tactical	-6.667	6.083	-1.096	.273	1.000
Gardening-Hockey	-12.667	6.083	-2.082	.037	.560
Tactical-Hockey	-6.000	6.083	-.986	.324	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.


Pairwise Comparisons of Glove
Power mesh
8.17
Gardening
20.83
Hockey
33.50
No glove
6.50
Running
14.50
Tactical
27.50
Each node shows the sample average rank of Glove.

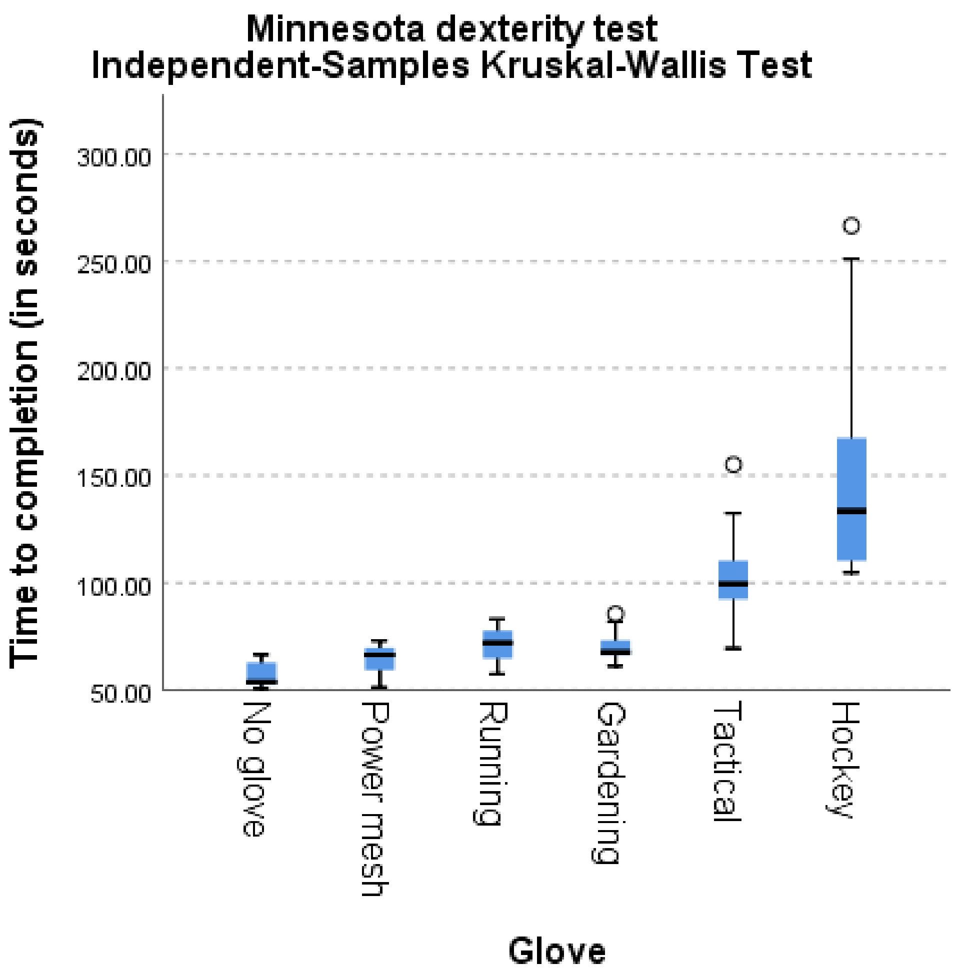

Continuous Field Information
O'Connor dexterity test - Time to completion
N = 36
Min = 112.95
Max = 1047.48
Mean = 317.201
Std. Dev. = 224.3946
Frequency
15
10
5
0
.00
200.00
400.00
600.00
800.00
1000.00
1200.00
O'Connor dexterity test - Time to completion (in s)

C.1.4 Minnesota dexterity test

Independent-Samples Kruskal-Wallis Test Summary

Total N	72
Test Statistic	54.374[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



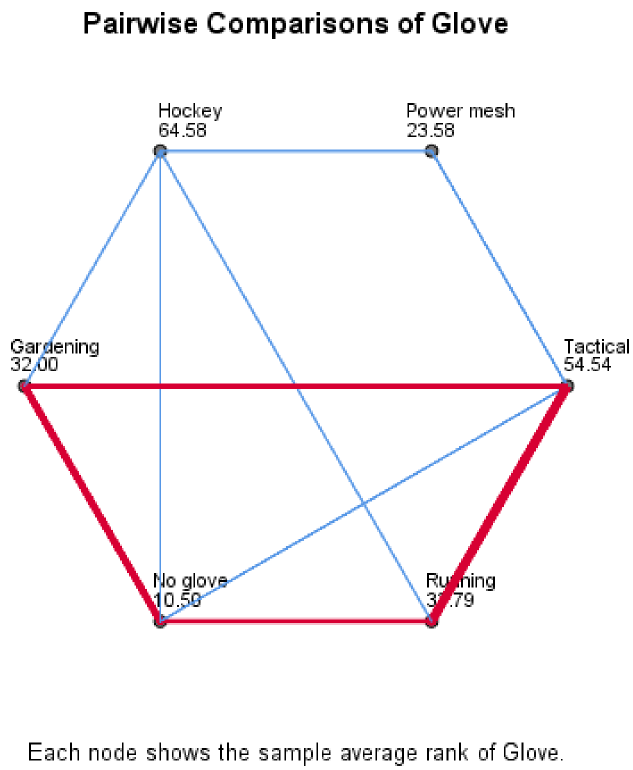
Pairwise Comparisons of Glove

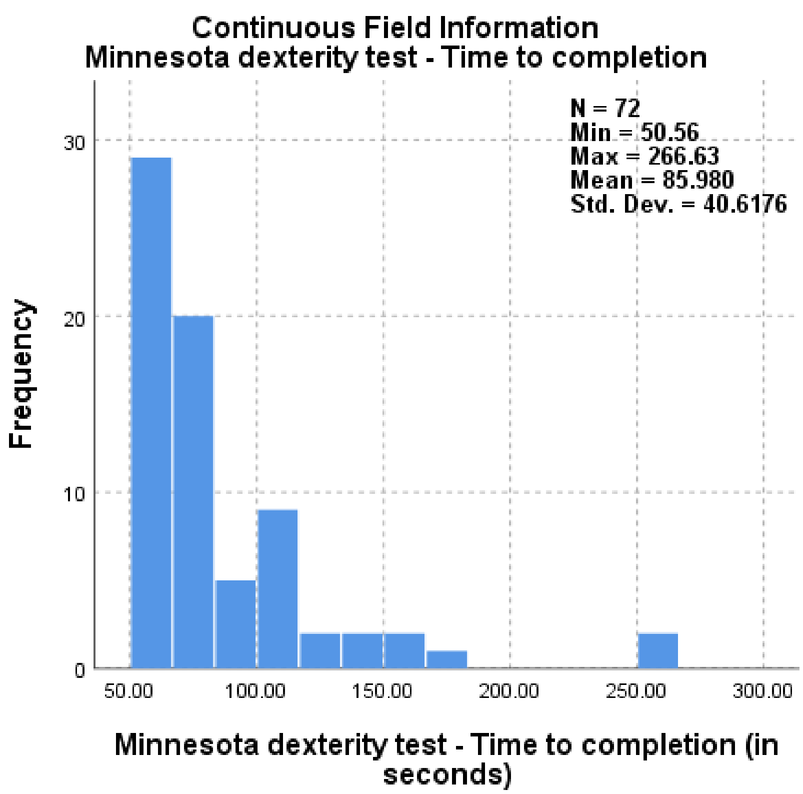
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
No glove-Power mesh	-13.083	8.544	-1.531	.126	1.000
No glove-Gardening	-21.500	8.544	-2.516	.012	.178
No glove-Running	-23.292	8.544	-2.726	.006	.096
No glove-Tactical	-44.042	8.544	-5.155	.000	.000
No glove-Hockey	-54.083	8.544	-6.330	.000	.000
Power mesh-Gardening	-8.417	8.544	-.985	.325	1.000
Power mesh-Running	-10.208	8.544	-1.195	.232	1.000
Power mesh-Tactical	-30.958	8.544	-3.623	.000	.004
Power mesh-Hockey	-41.000	8.544	-4.799	.000	.000
Gardening-Running	1.792	8.544	.210	.834	1.000
Gardening-Tactical	-22.542	8.544	-2.638	.008	.125
Gardening-Hockey	-32.583	8.544	-3.814	.000	.002
Running-Tactical	-20.750	8.544	-2.429	.015	.227
Running-Hockey	-30.792	8.544	-3.604	.000	.005
Tactical-Hockey	-10.042	8.544	-1.175	.240	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.


Pairwise Comparisons of Glove
Hockey
64.58
Power mesh
23.58
Gardening
32.00
Tactical
54.54
No glove
10.50
Running
Each node shows the sample average rank of Glove.

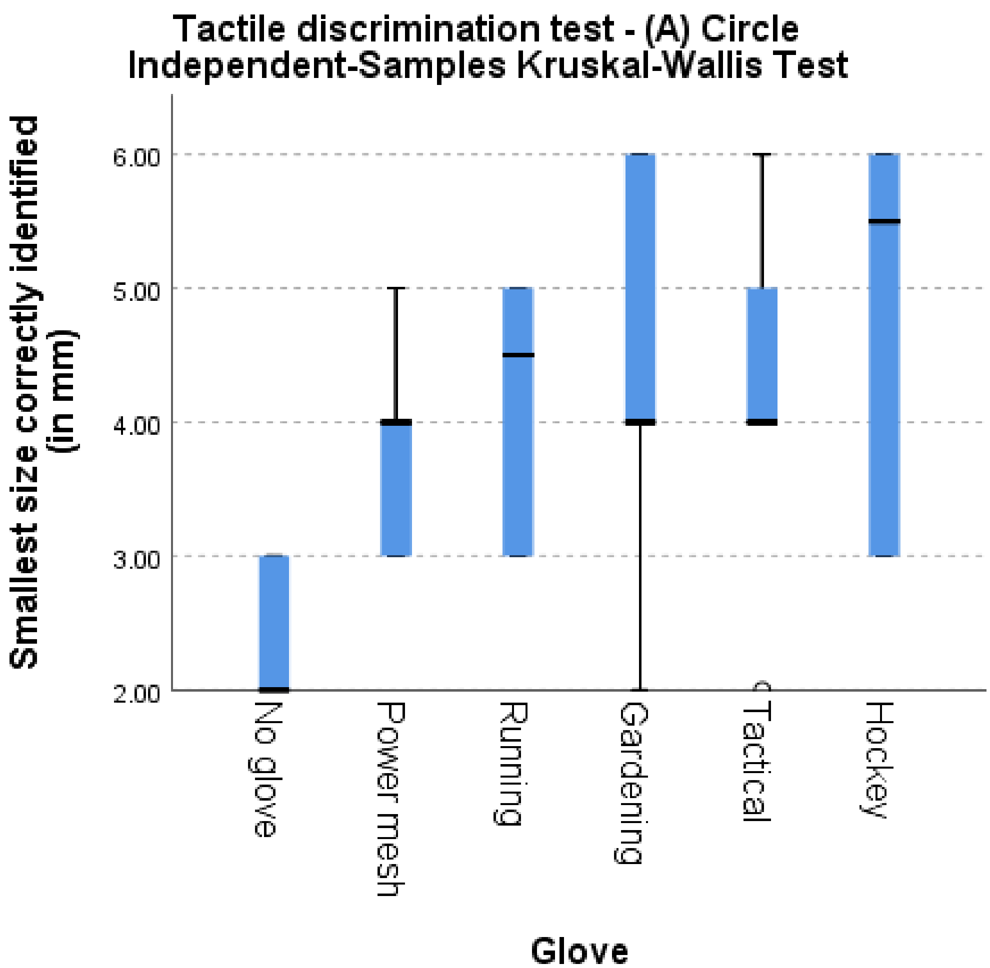

Continuous Field Information
Minnesota dexterity test - Time to completion
N = 72
Min = 50.56
Max = 266.63
Mean = 85.980
Std. Dev. = 40.6176
Frequency
0
10
20
30
50.00
100.00
150.00
200.00
250.00
300.00
Minnesota dexterity test - Time to completion (in seconds)

C.1.5 Tactile discrimination test - (A) Circle

Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	12.346[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.030

a. The test statistic is adjusted for ties.



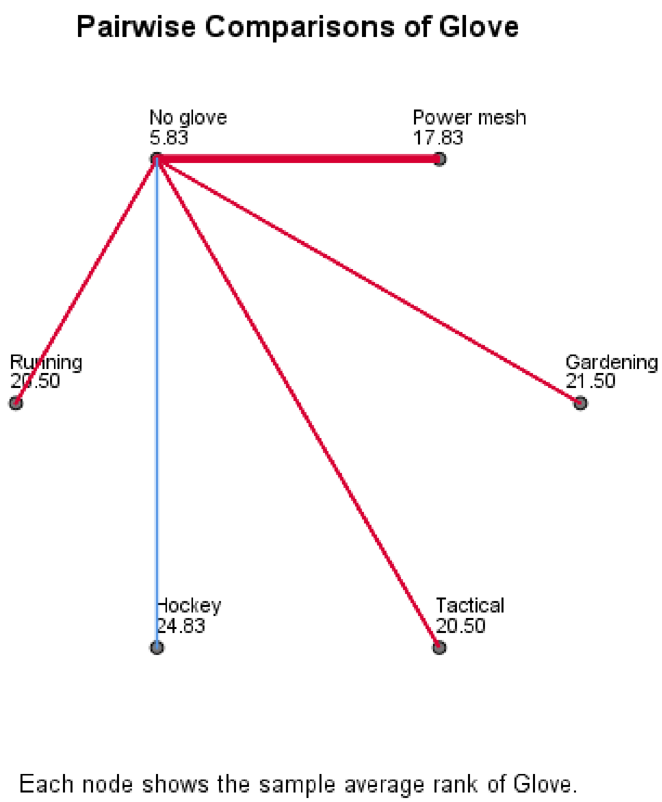
Pairwise Comparisons of Glove

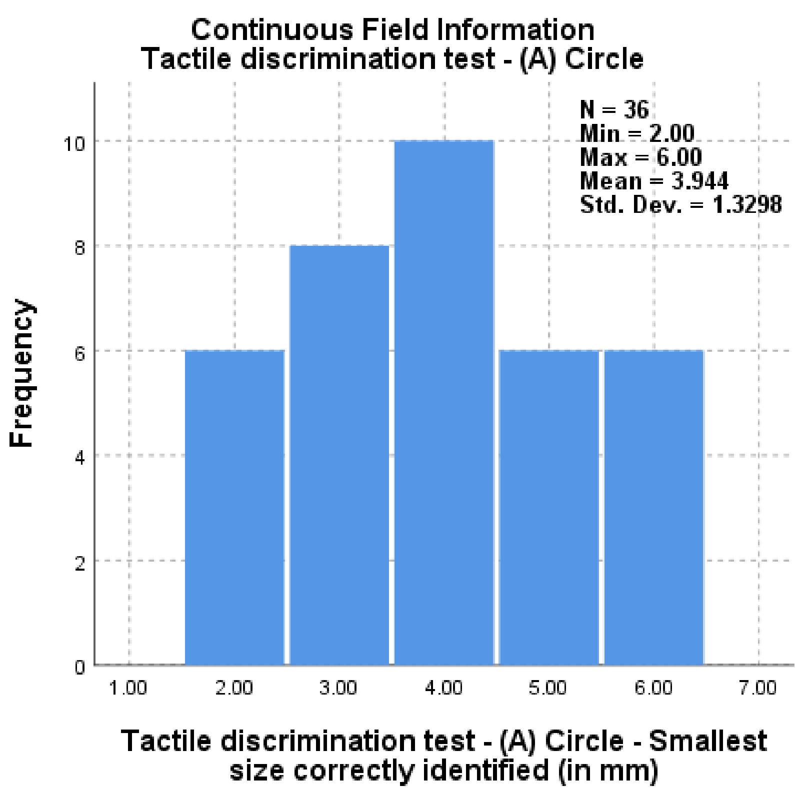
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
No glove-Power mesh	-12.000	5.943	-2.019	.043	.652
No glove-Running	-14.667	5.943	-2.468	.014	.204
No glove-Tactical	-14.667	5.943	-2.468	.014	.204
No glove-Gardening	-15.667	5.943	-2.636	.008	.126
No glove-Hockey	-19.000	5.943	-3.197	.001	.021
Power mesh-Running	-2.667	5.943	-.449	.654	1.000
Power mesh-Tactical	-2.667	5.943	-.449	.654	1.000
Power mesh-Gardening	-3.667	5.943	-.617	.537	1.000
Power mesh-Hockey	-7.000	5.943	-1.178	.239	1.000
Running-Tactical	.000	5.943	.000	1.000	1.000
Running-Gardening	-1.000	5.943	-.168	.866	1.000
Running-Hockey	-4.333	5.943	-.729	.466	1.000
Tactical-Gardening	1.000	5.943	.168	.866	1.000
Tactical-Hockey	-4.333	5.943	-.729	.466	1.000
Gardening-Hockey	-3.333	5.943	-.561	.575	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.


Pairwise Comparisons of Glove
No glove
5.83
Power mesh
17.83
Running
20.50
Gardening
21.50
Hockey
24.83
Tactical
20.50
Each node shows the sample average rank of Glove.


Continuous Field Information
Tactile discrimination test - (A) Circle
N = 36
Min = 2.00
Max = 6.00
Mean = 3.944
Std. Dev. = 1.3298
Frequency
0
2
4
6
8
10
1.00
2.00
3.00
4.00
5.00
6.00
7.00
Tactile discrimination test - (A) Circle - Smallest size correctly identified (in mm)

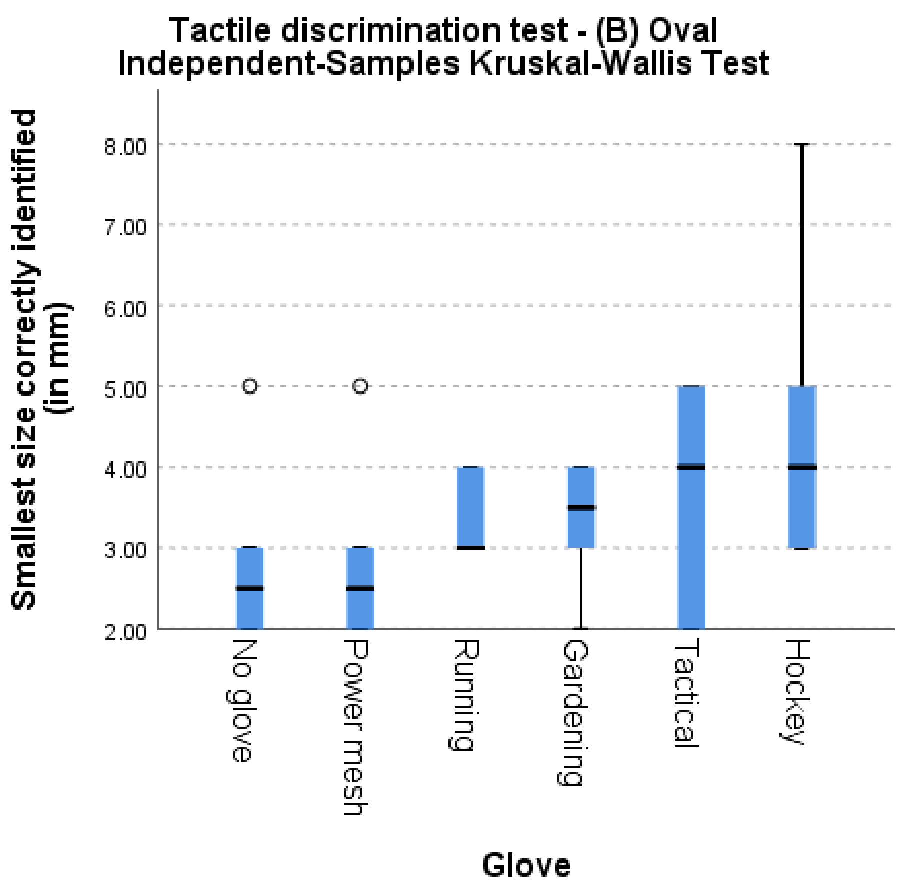
C.1.6 Tactile discrimination test - (B) Oval

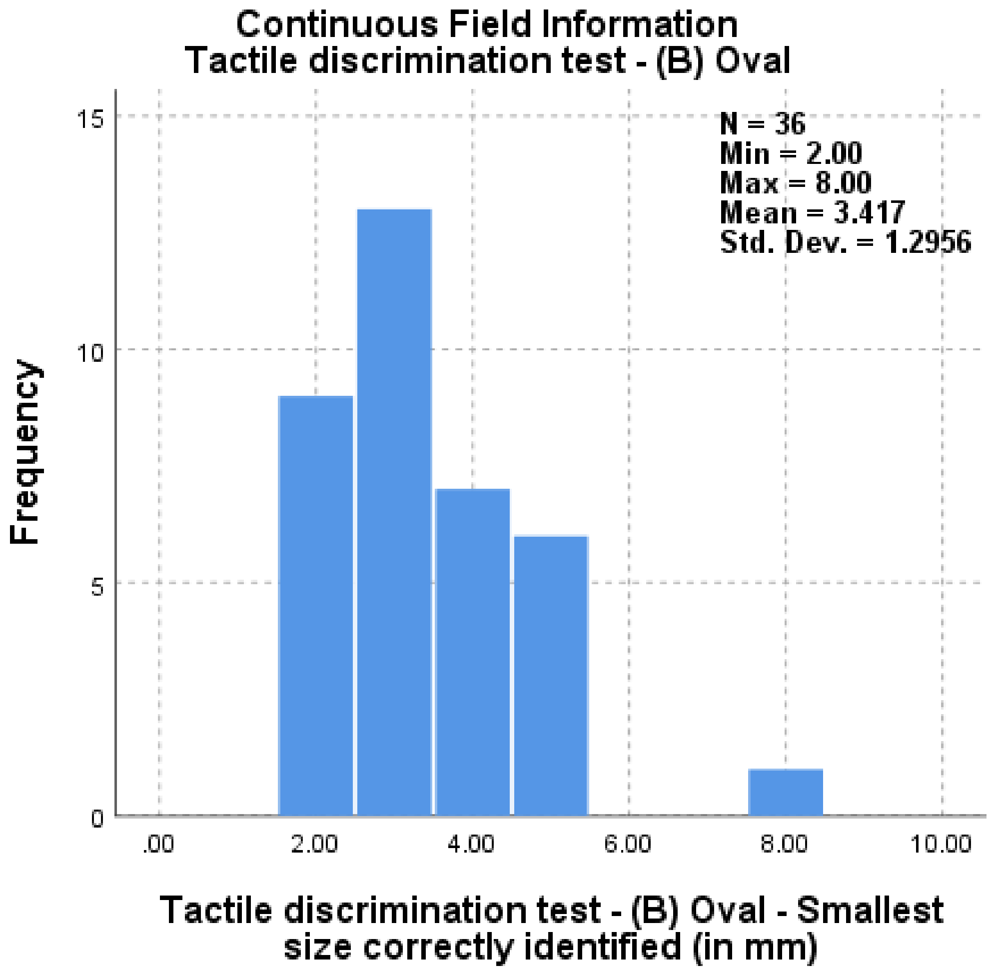
Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	6.041[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.302

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





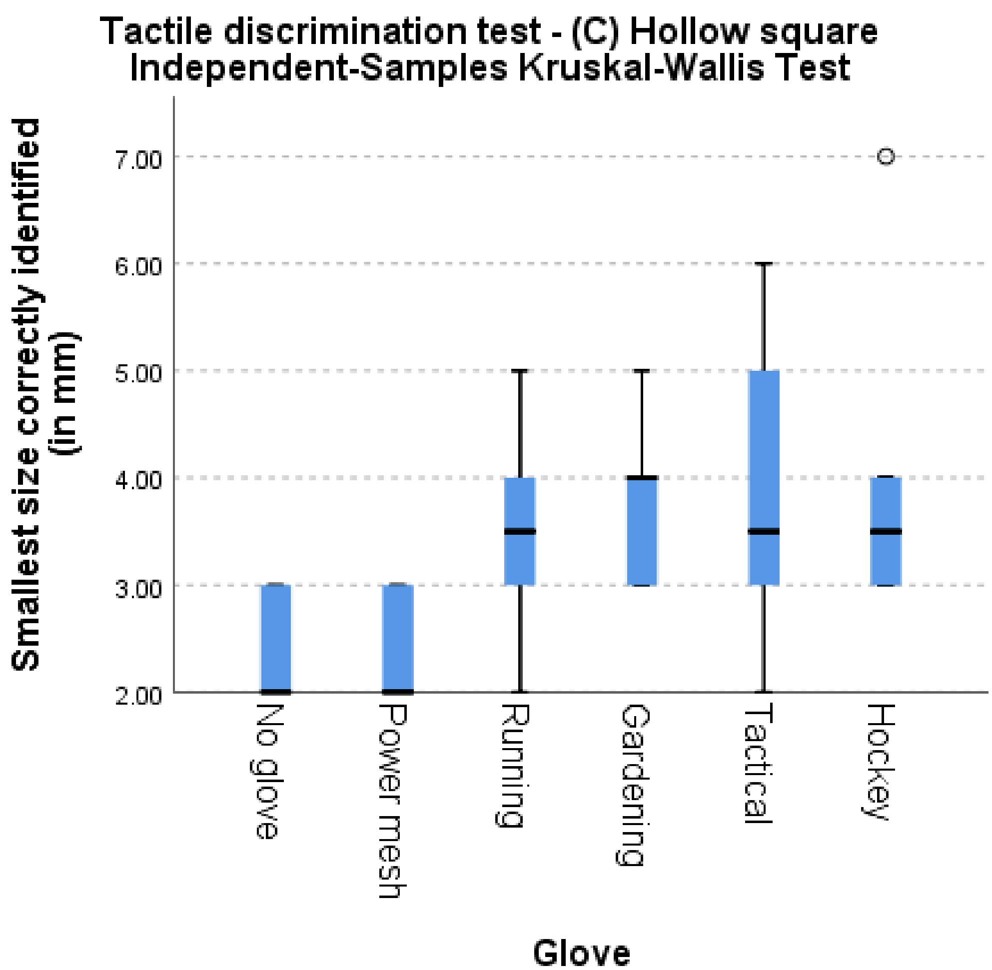
Continuous Field Information
Tactile discrimination test - (B) Oval
N = 36
Min = 2.00
Max = 8.00
Mean = 3.417
Std. Dev. = 1.2956
Frequency
15
10
5
0
.00
2.00
4.00
6.00
8.00
10.00
Tactile discrimination test - (B) Oval - Smallest size correctly identified (in mm)

C.1.7 Tactile discrimination test - (C) Hollow square

Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	15.236[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.009

a. The test statistic is adjusted for ties.



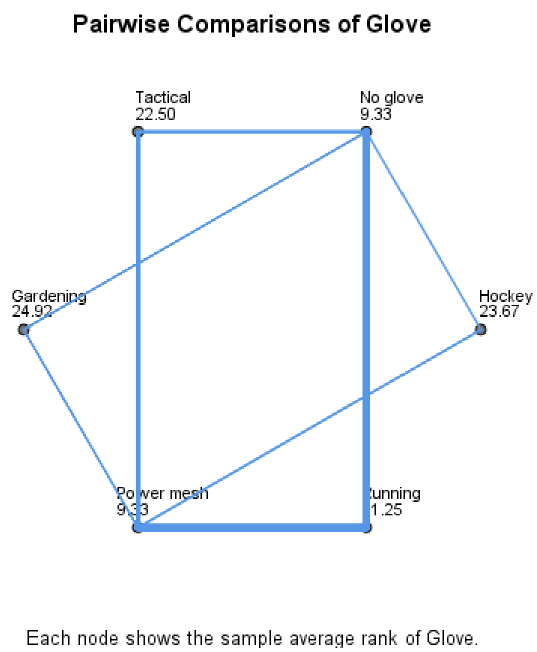
Pairwise Comparisons of Glove

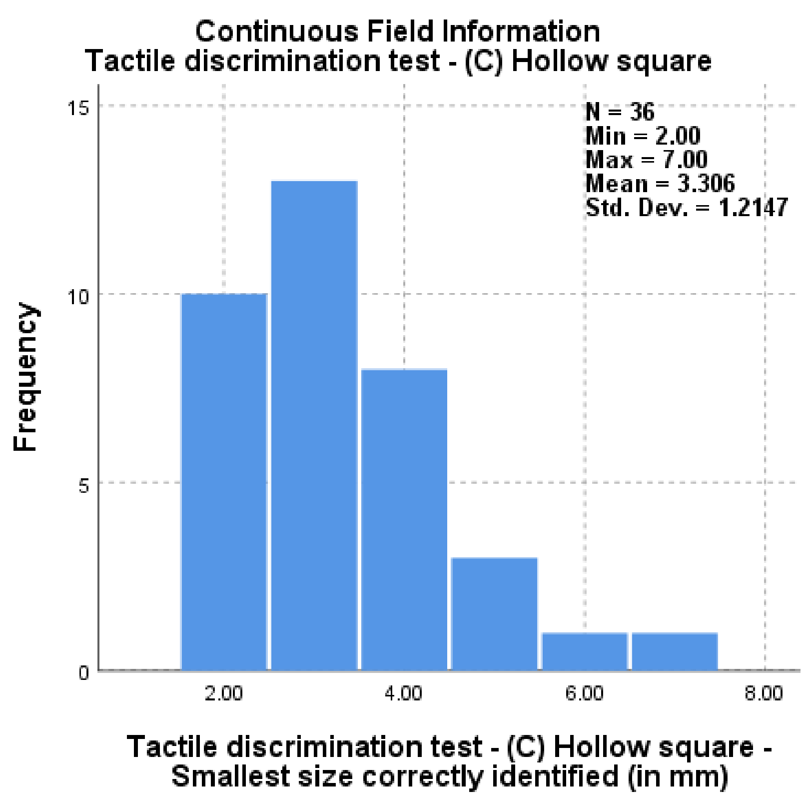
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
No glove-Power mesh	.000	5.836	.000	1.000	1.000
No glove-Running	-11.917	5.836	-2.042	.041	.618
No glove-Tactical	-13.167	5.836	-2.256	.024	.361
No glove-Hockey	-14.333	5.836	-2.456	.014	.211
No glove-Gardening	-15.583	5.836	-2.670	.008	.114
Power mesh-Running	-11.917	5.836	-2.042	.041	.618
Power mesh-Tactical	-13.167	5.836	-2.256	.024	.361
Power mesh-Hockey	-14.333	5.836	-2.456	.014	.211
Power mesh-Gardening	-15.583	5.836	-2.670	.008	.114
Running-Tactical	-1.250	5.836	-.214	.830	1.000
Running-Hockey	-2.417	5.836	-.414	.679	1.000
Running-Gardening	-3.667	5.836	-.628	.530	1.000
Tactical-Hockey	-1.167	5.836	-.200	.842	1.000
Tactical-Gardening	2.417	5.836	.414	.679	1.000
Hockey-Gardening	1.250	5.836	.214	.830	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.


Pairwise Comparisons of Glove
Tactical
22.50
No glove
9.33
Gardening
24.92
Hockey
23.67
Power mesh
9.33
Running
1.25
Each node shows the sample average rank of Glove.


Continuous Field Information
Tactile discrimination test - (C) Hollow square
N = 36
Min = 2.00
Max = 7.00
Mean = 3.306
Std. Dev. = 1.2147
Frequency
15
10
5
0
2.00
4.00
6.00
8.00
Tactile discrimination test - (C) Hollow square - Smallest size correctly identified (in mm)

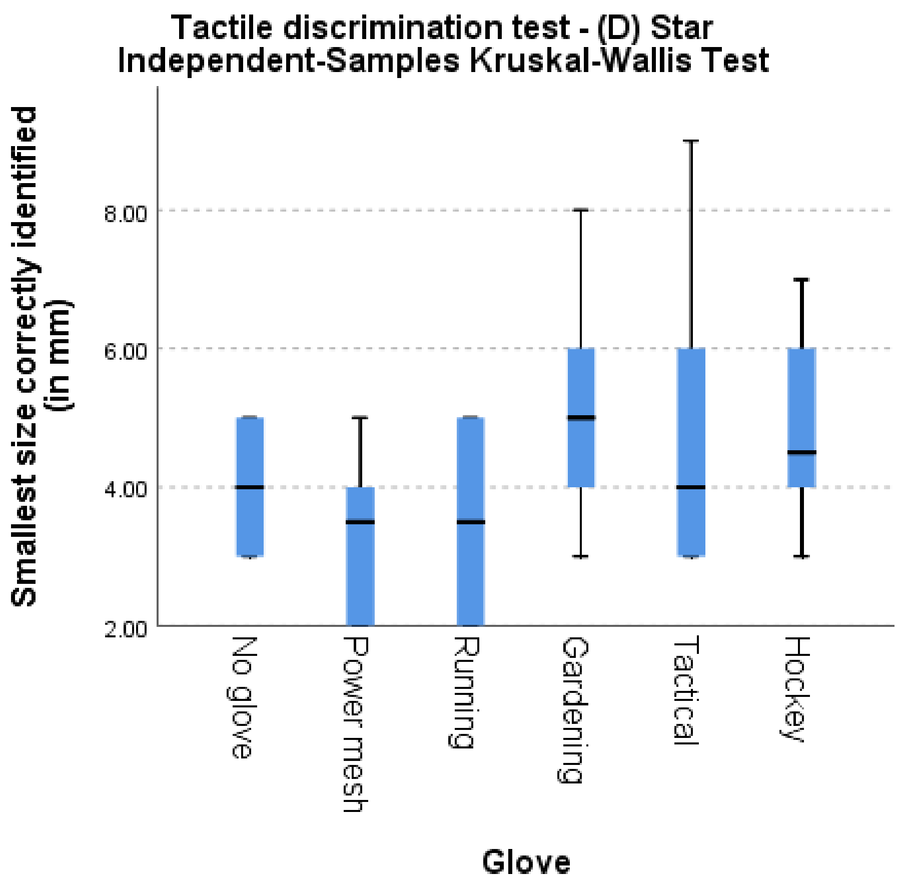
C.1.8 Tactile discrimination test - (D) Star

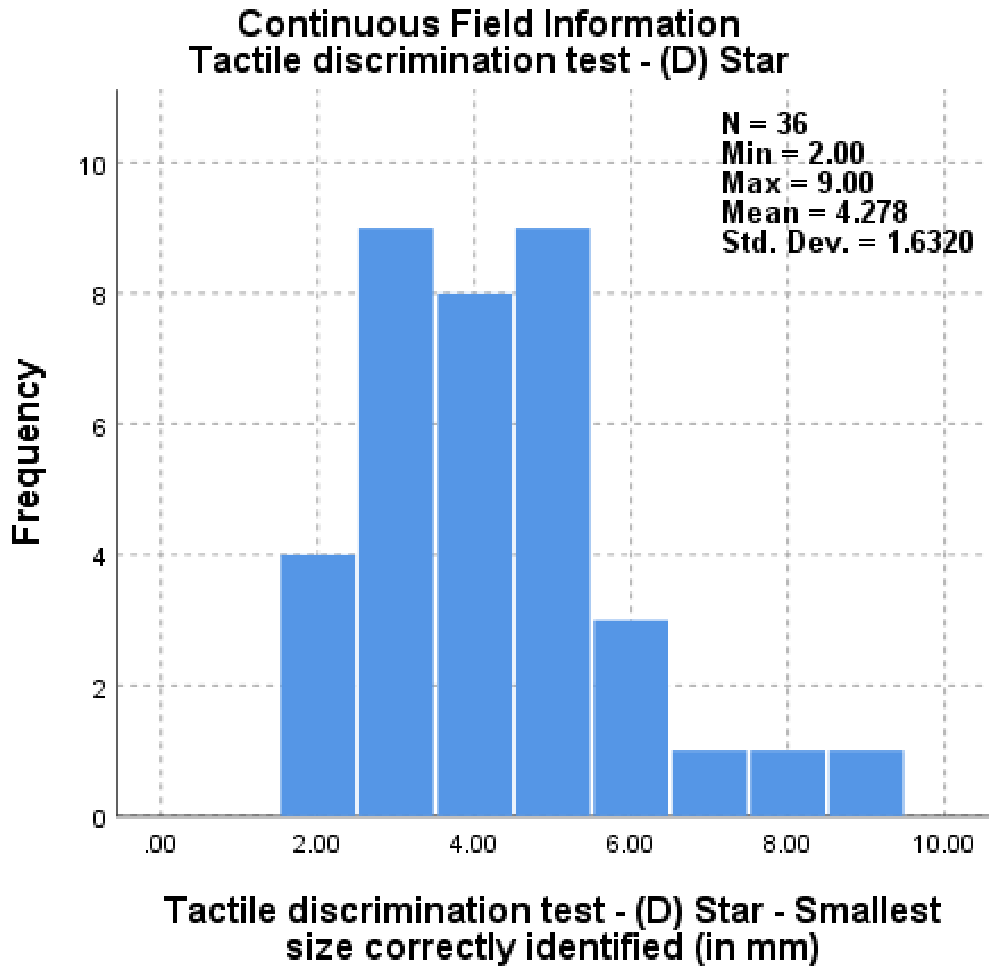
Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	6.285[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.279

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





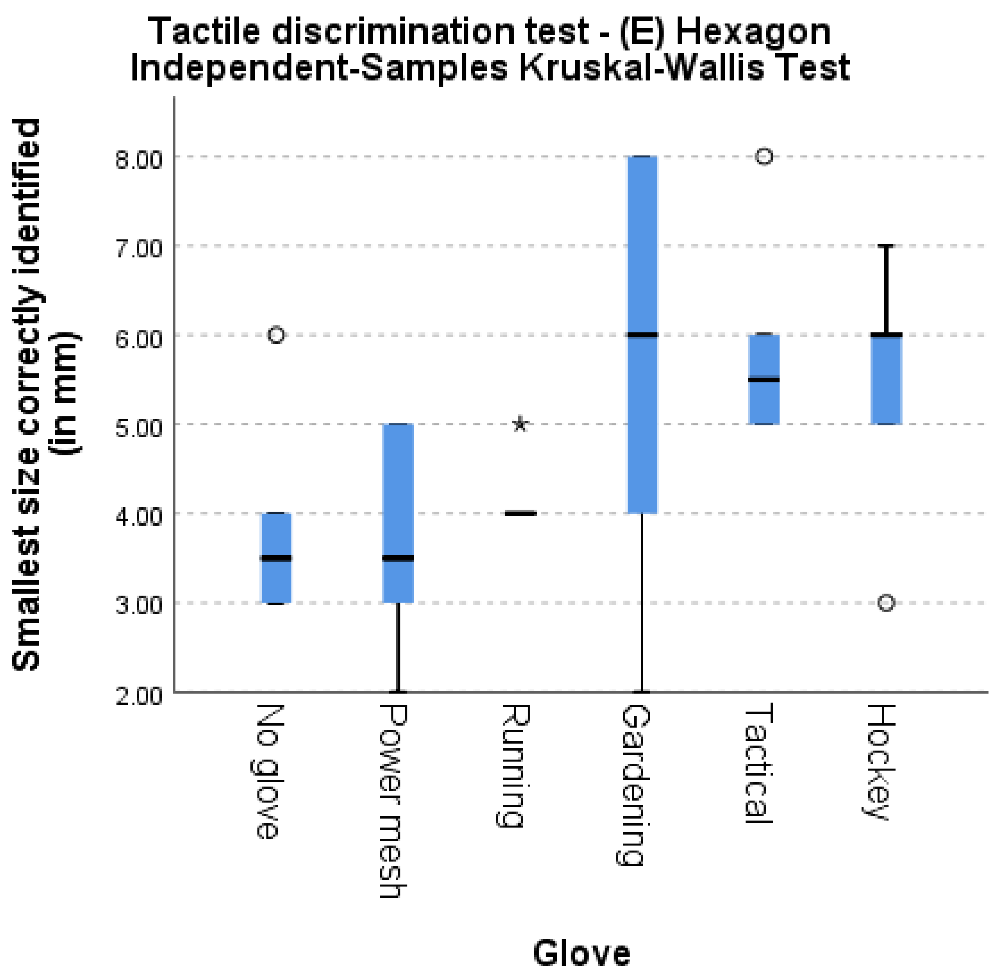
Continuous Field Information
Tactile discrimination test - (D) Star
N = 36
Min = 2.00
Max = 9.00
Mean = 4.278
Std. Dev. = 1.6320
Frequency
0
2
4
6
8
10
.00
2.00
4.00
6.00
8.00
10.00
Tactile discrimination test - (D) Star - Smallest size correctly identified (in mm)

C.1.9 Tactile discrimination test - (E) Hexagon

Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	12.712[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.026

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

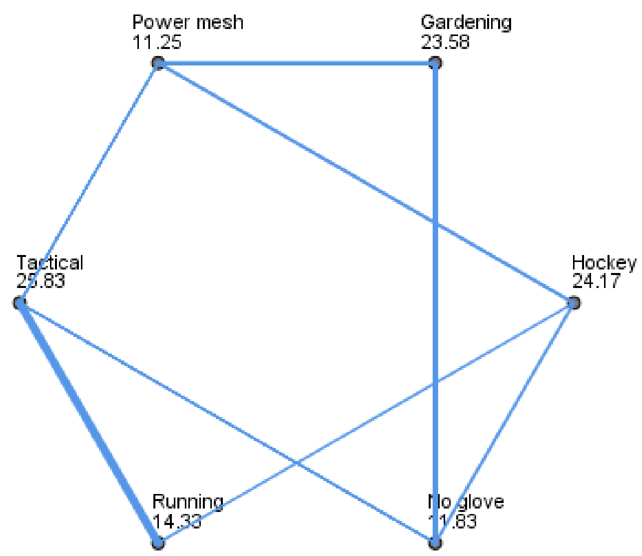
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Power mesh-No glove	.583	5.964	.098	.922	1.000
Power mesh-Running	-3.083	5.964	-.517	.605	1.000
Power mesh-Gardening	-12.333	5.964	-2.068	.039	.580
Power mesh-Hockey	-12.917	5.964	-2.166	.030	.455
Power mesh-Tactical	-14.583	5.964	-2.445	.014	.217
No glove-Running	-2.500	5.964	-.419	.675	1.000
No glove-Gardening	-11.750	5.964	-1.970	.049	.732
No glove-Hockey	-12.333	5.964	-2.068	.039	.580
No glove-Tactical	-14.000	5.964	-2.347	.019	.284
Running-Gardening	-9.250	5.964	-1.551	.121	1.000
Running-Hockey	-9.833	5.964	-1.649	.099	1.000
Running-Tactical	-11.500	5.964	-1.928	.054	.807
Gardening-Hockey	-.583	5.964	-.098	.922	1.000
Gardening-Tactical	-2.250	5.964	-.377	.706	1.000
Hockey-Tactical	1.667	5.964	.279	.780	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

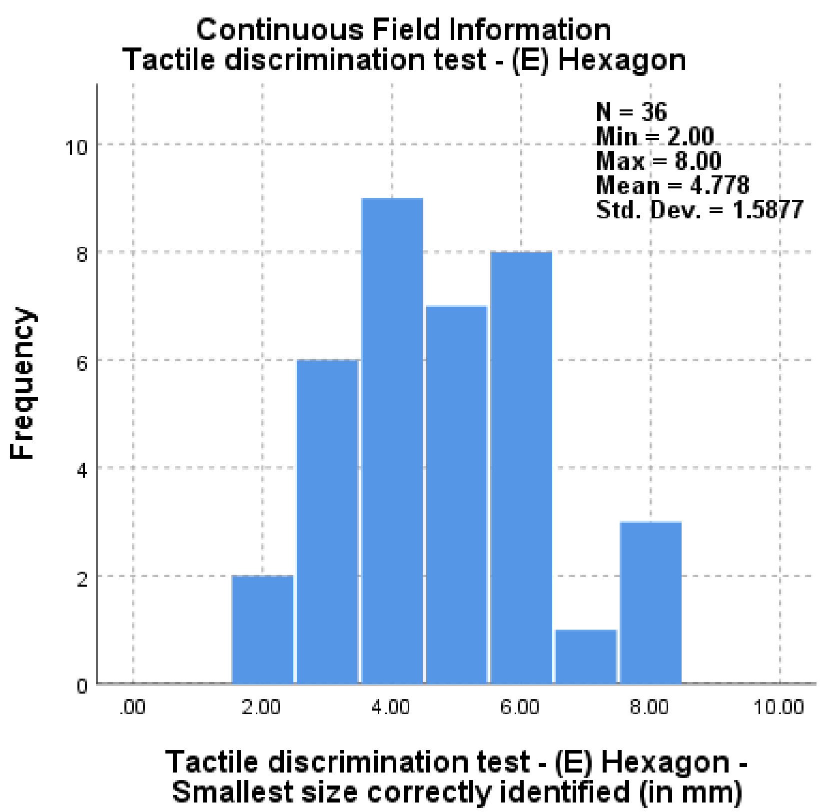
Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



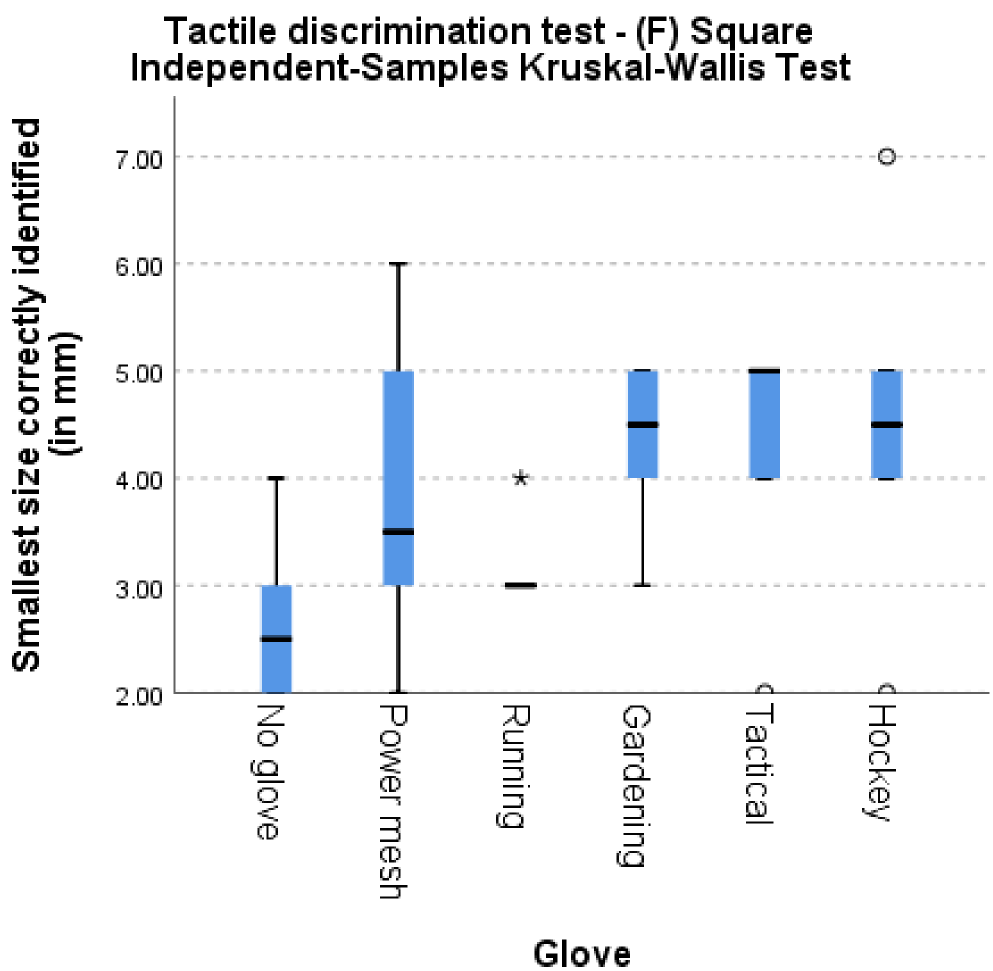
Each node shows the sample average rank of Glove.



C.1.10 Tactile discrimination test - (F) Square

Independent-Samples Kruskal-Wallis Test Summary	
Total N	36
Test Statistic	11.173[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.048

a. The test statistic is adjusted for ties.



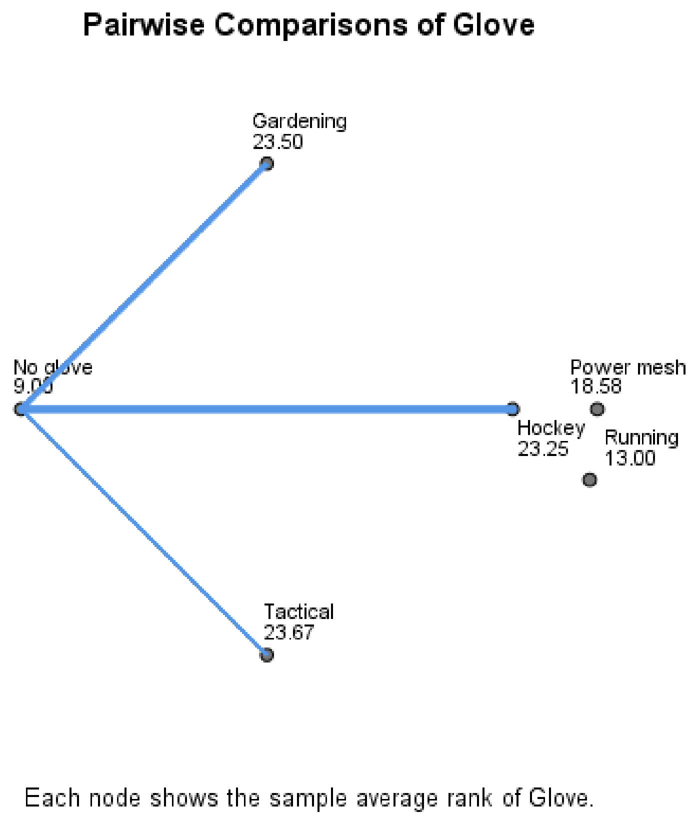
Pairwise Comparisons of Glove

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
No glove-Running	-4.000	5.904	-.677	.498	1.000
No glove-Power mesh	-9.583	5.904	-1.623	.105	1.000
No glove-Hockey	-14.250	5.904	-2.413	.016	.237
No glove-Gardening	-14.500	5.904	-2.456	.014	.211
No glove-Tactical	-14.667	5.904	-2.484	.013	.195
Running-Power mesh	5.583	5.904	.946	.344	1.000
Running-Hockey	-10.250	5.904	-1.736	.083	1.000
Running-Gardening	-10.500	5.904	-1.778	.075	1.000
Running-Tactical	-10.667	5.904	-1.807	.071	1.000
Power mesh-Hockey	-4.667	5.904	-.790	.429	1.000
Power mesh-Gardening	-4.917	5.904	-.833	.405	1.000
Power mesh-Tactical	-5.083	5.904	-.861	.389	1.000
Hockey-Gardening	.250	5.904	.042	.966	1.000
Hockey-Tactical	.417	5.904	.071	.944	1.000
Gardening-Tactical	-.167	5.904	-.028	.977	1.000

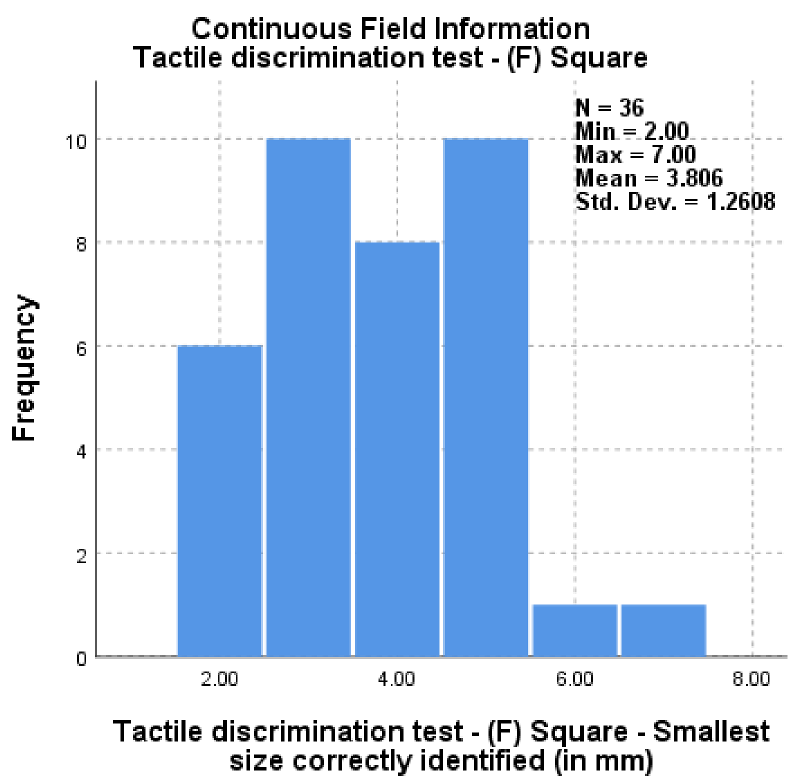
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



Pairwise Comparisons of Glove
Gardening
23.50
No glove
9.00
Power mesh
18.58
Hockey
23.25
Running
13.00
Tactical
23.67
Each node shows the sample average rank of Glove.



Continuous Field Information
Tactile discrimination test - (F) Square
N = 36
Min = 2.00
Max = 7.00
Mean = 3.806
Std. Dev. = 1.2608
Frequency
0
2
4
6
8
10
2.00
4.00
6.00
8.00
Tactile discrimination test - (F) Square - Smallest size correctly identified (in mm)

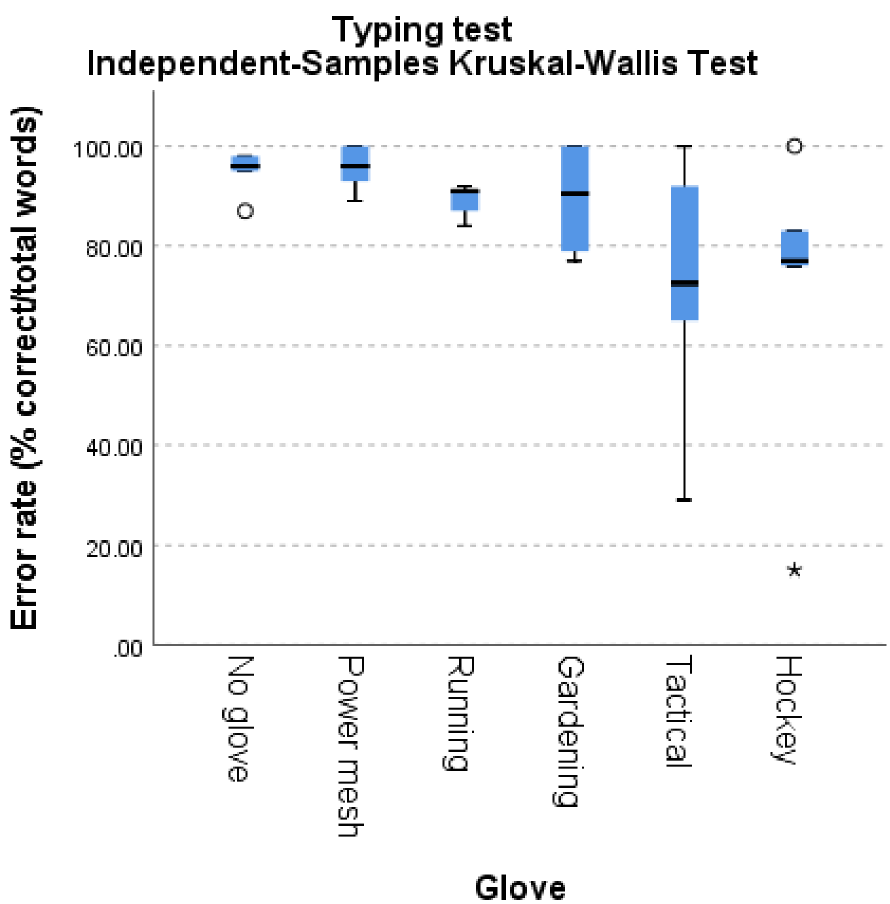
C.2 Error rate data

C.2.1 Typing test - Error rate

Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	11.386[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.044

a. The test statistic is adjusted for ties.



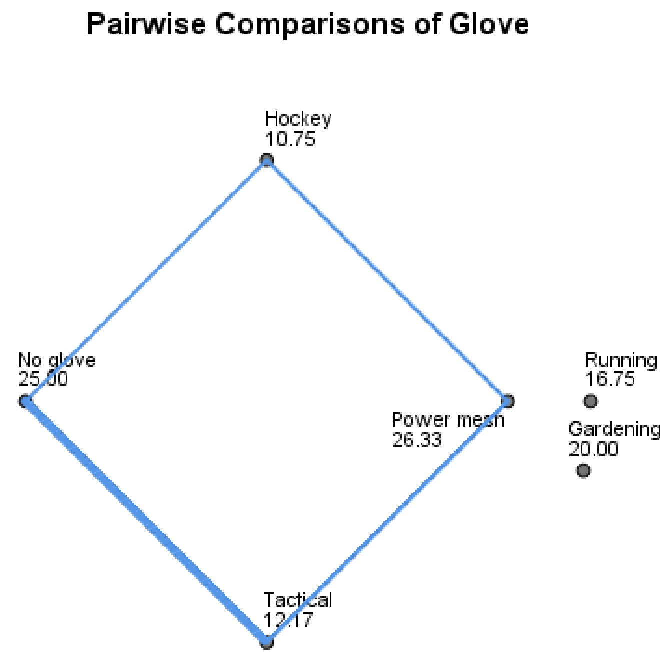
Pairwise Comparisons of Glove

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Hockey-Tactical	1.417	6.060	.234	.815	1.000
Hockey-Running	6.000	6.060	.990	.322	1.000
Hockey-Gardening	9.250	6.060	1.526	.127	1.000
Hockey-No glove	14.250	6.060	2.351	.019	.281
Hockey-Power mesh	15.583	6.060	2.571	.010	.152
Tactical-Running	4.583	6.060	.756	.449	1.000
Tactical-Gardening	7.833	6.060	1.293	.196	1.000
Tactical-No glove	12.833	6.060	2.118	.034	.513
Tactical-Power mesh	14.167	6.060	2.338	.019	.291
Running-Gardening	-3.250	6.060	-.536	.592	1.000
Running-No glove	8.250	6.060	1.361	.173	1.000
Running-Power mesh	9.583	6.060	1.581	.114	1.000
Gardening-No glove	5.000	6.060	.825	.409	1.000
Gardening-Power mesh	6.333	6.060	1.045	.296	1.000
No glove-Power mesh	-1.333	6.060	-.220	.826	1.000

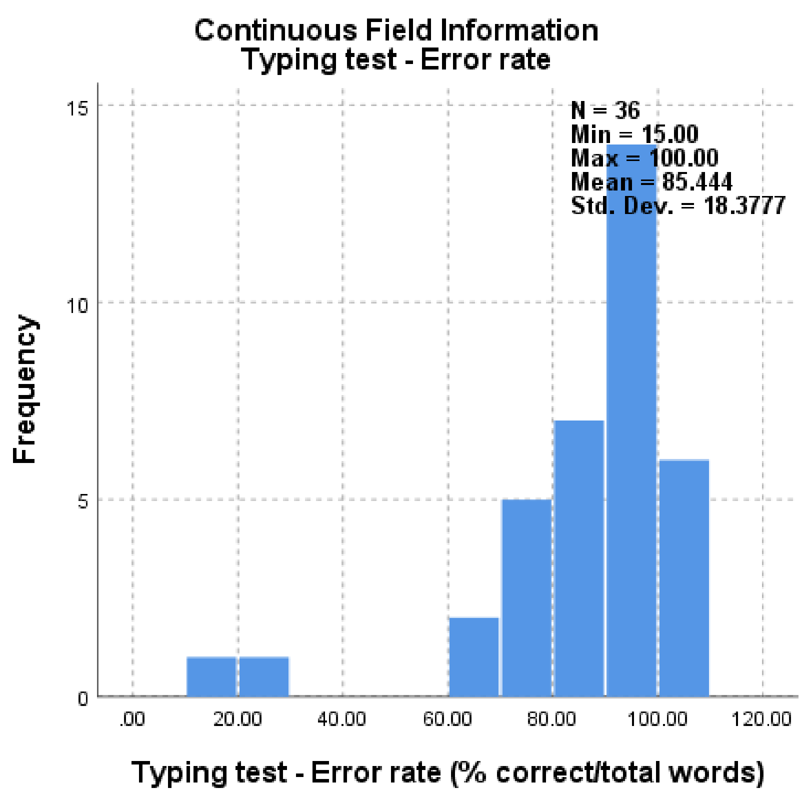
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



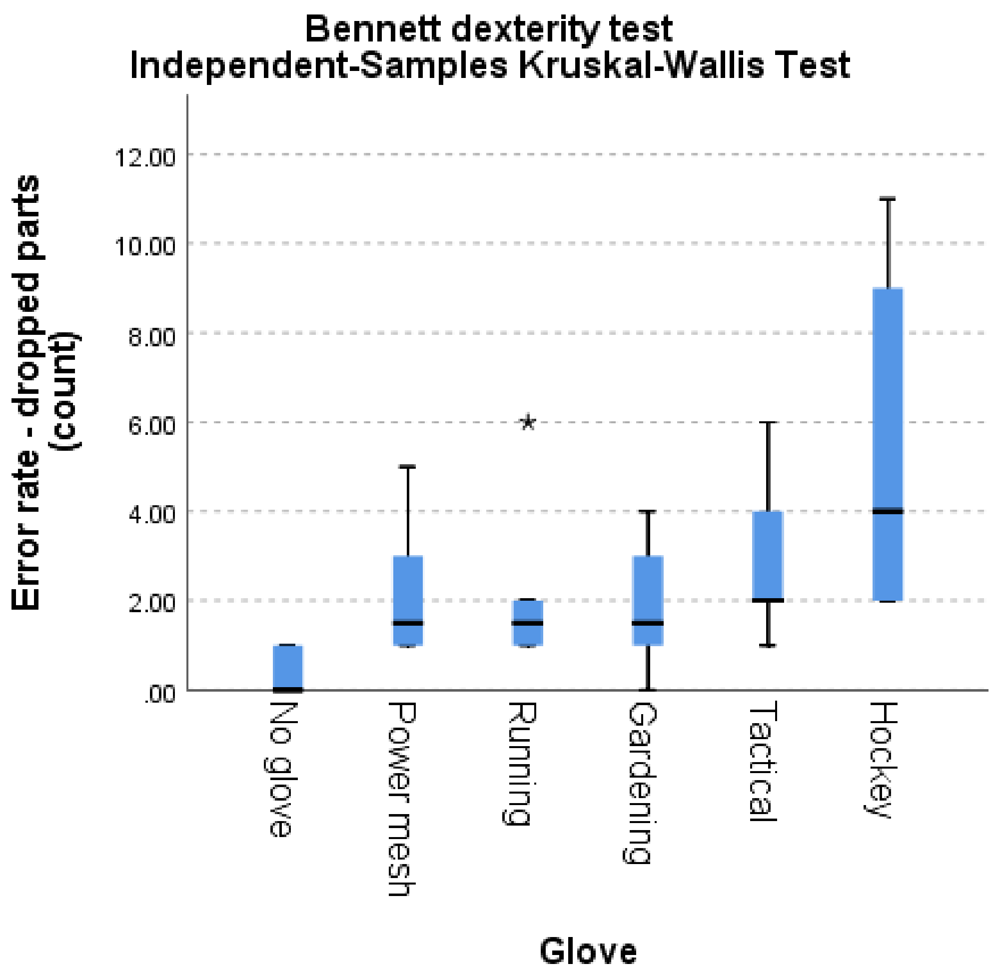
Each node shows the sample average rank of Glove.



C.2.2 Bennett dexterity test - Error rate

Independent-Samples Kruskal-Wallis Test Summary	
Total N	36
Test Statistic	16.305[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.006

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

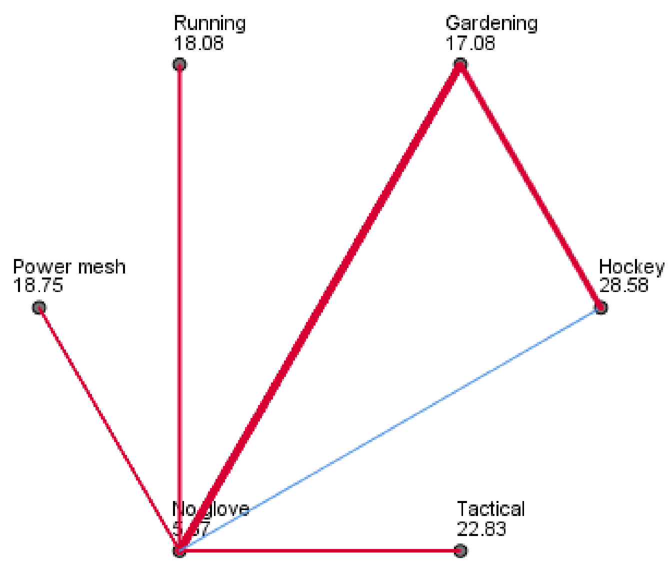
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
No glove-Gardening	-11.417	5.937	-1.923	.054	.817
No glove-Running	-12.417	5.937	-2.091	.037	.548
No glove-Power mesh	-13.083	5.937	-2.204	.028	.413
No glove-Tactical	-17.167	5.937	-2.891	.004	.058
No glove-Hockey	-22.917	5.937	-3.860	.000	.002
Gardening-Running	1.000	5.937	.168	.866	1.000
Gardening-Power mesh	1.667	5.937	.281	.779	1.000
Gardening-Tactical	-5.750	5.937	-.968	.333	1.000
Gardening-Hockey	-11.500	5.937	-1.937	.053	.791
Running-Power mesh	.667	5.937	.112	.911	1.000
Running-Tactical	-4.750	5.937	-.800	.424	1.000
Running-Hockey	-10.500	5.937	-1.768	.077	1.000
Power mesh-Tactical	-4.083	5.937	-.688	.492	1.000
Power mesh-Hockey	-9.833	5.937	-1.656	.098	1.000
Tactical-Hockey	-5.750	5.937	-.968	.333	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

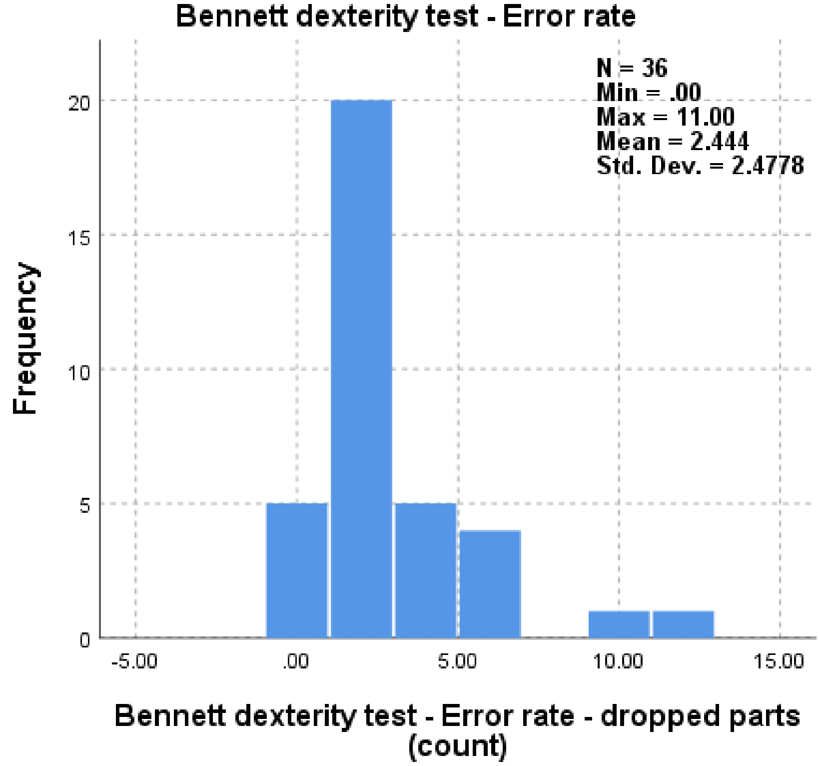
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.

Continuous Field Information
Bennett dexterity test - Error rate

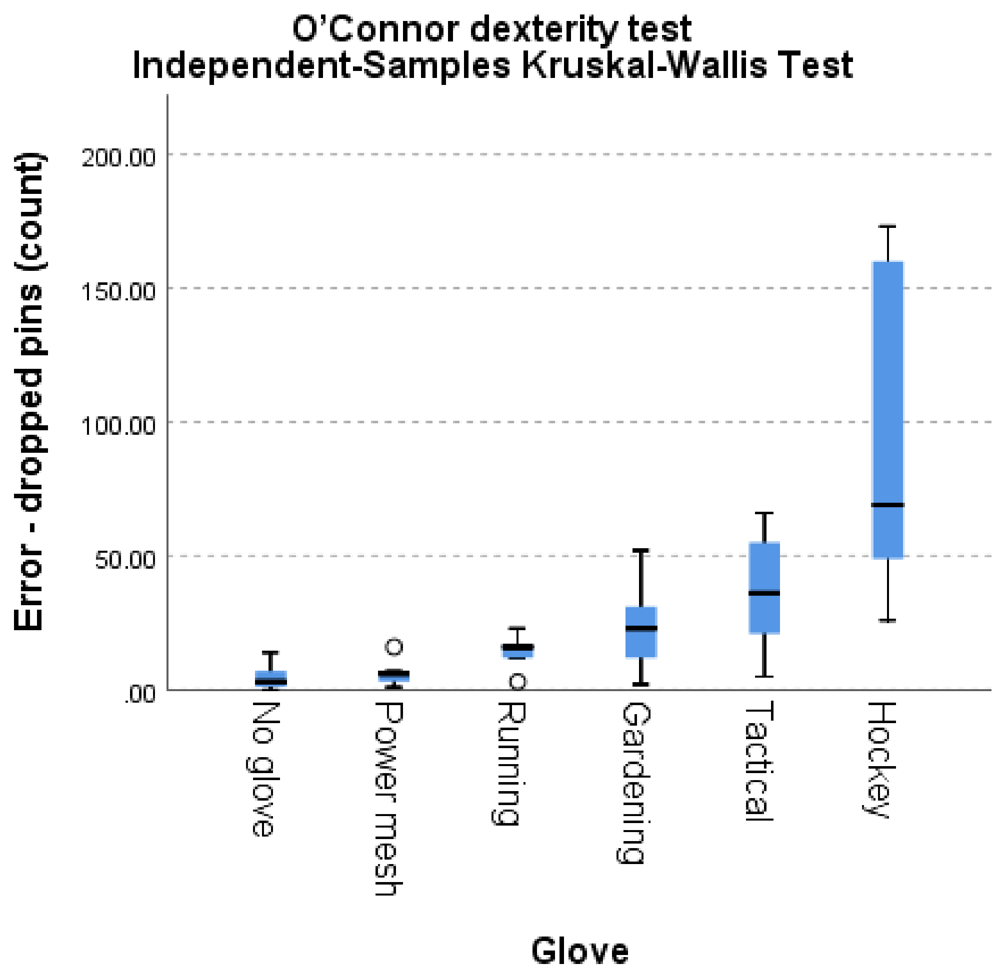


C.2.3 O'Connor dexterity test - Error rate

Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	22.300[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

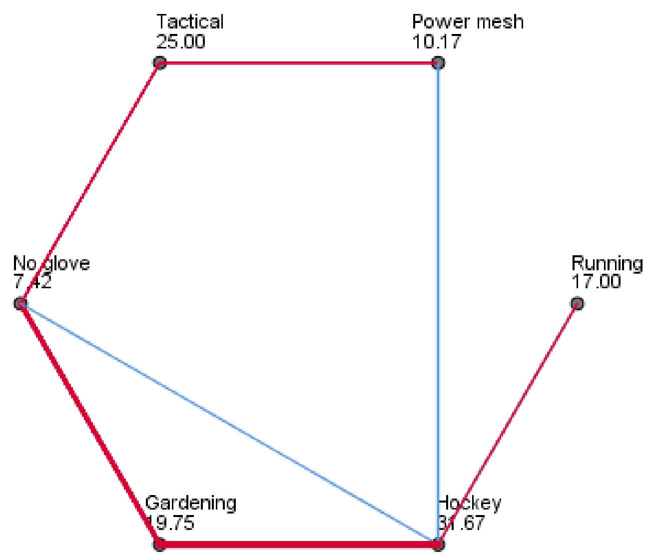
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
No glove-Power mesh	-2.750	6.076	-.453	.651	1.000
No glove-Running	-9.583	6.076	-1.577	.115	1.000
No glove-Gardening	-12.333	6.076	-2.030	.042	.636
No glove-Tactical	-17.583	6.076	-2.894	.004	.057
No glove-Hockey	-24.250	6.076	-3.991	.000	.001
Power mesh-Running	-6.833	6.076	-1.125	.261	1.000
Power mesh-Gardening	-9.583	6.076	-1.577	.115	1.000
Power mesh-Tactical	-14.833	6.076	-2.441	.015	.220
Power mesh-Hockey	-21.500	6.076	-3.538	.000	.006
Running-Gardening	-2.750	6.076	-.453	.651	1.000
Running-Tactical	-8.000	6.076	-1.317	.188	1.000
Running-Hockey	-14.667	6.076	-2.414	.016	.237
Gardening-Tactical	-5.250	6.076	-.864	.388	1.000
Gardening-Hockey	-11.917	6.076	-1.961	.050	.748
Tactical-Hockey	-6.667	6.076	-1.097	.273	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

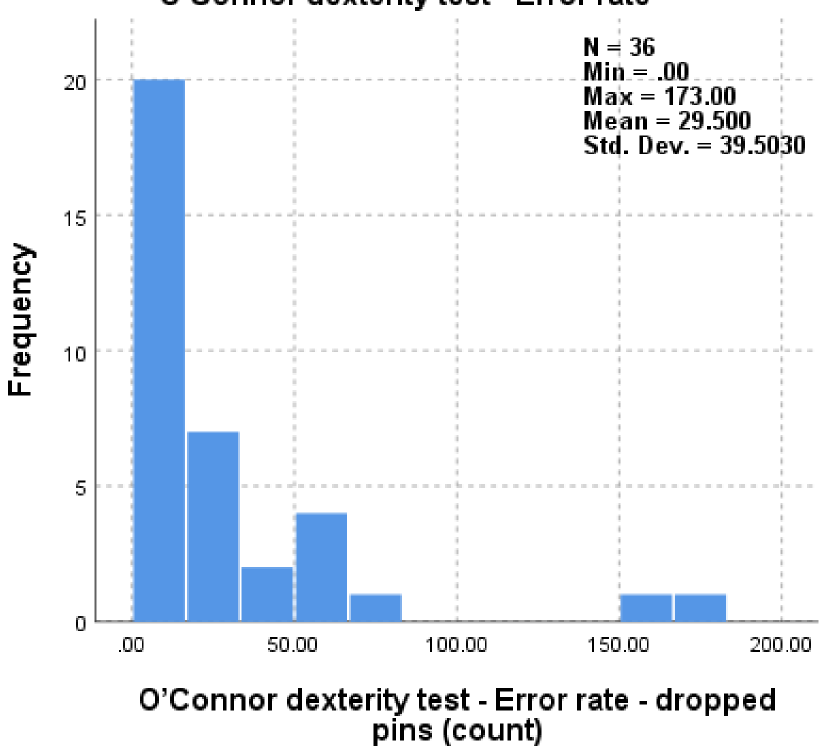
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.

Continuous Field Information
O'Connor dexterity test - Error rate



C.3 Task NASA-TLX data

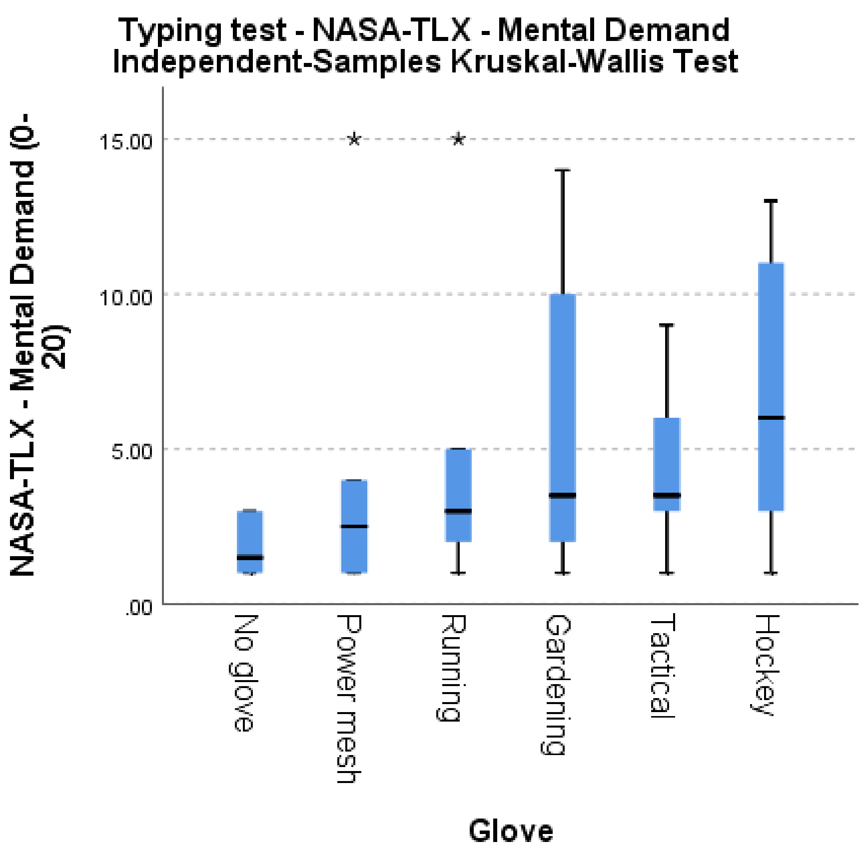
C.3.1 Typing test - NASA-TLX - Mental Demand

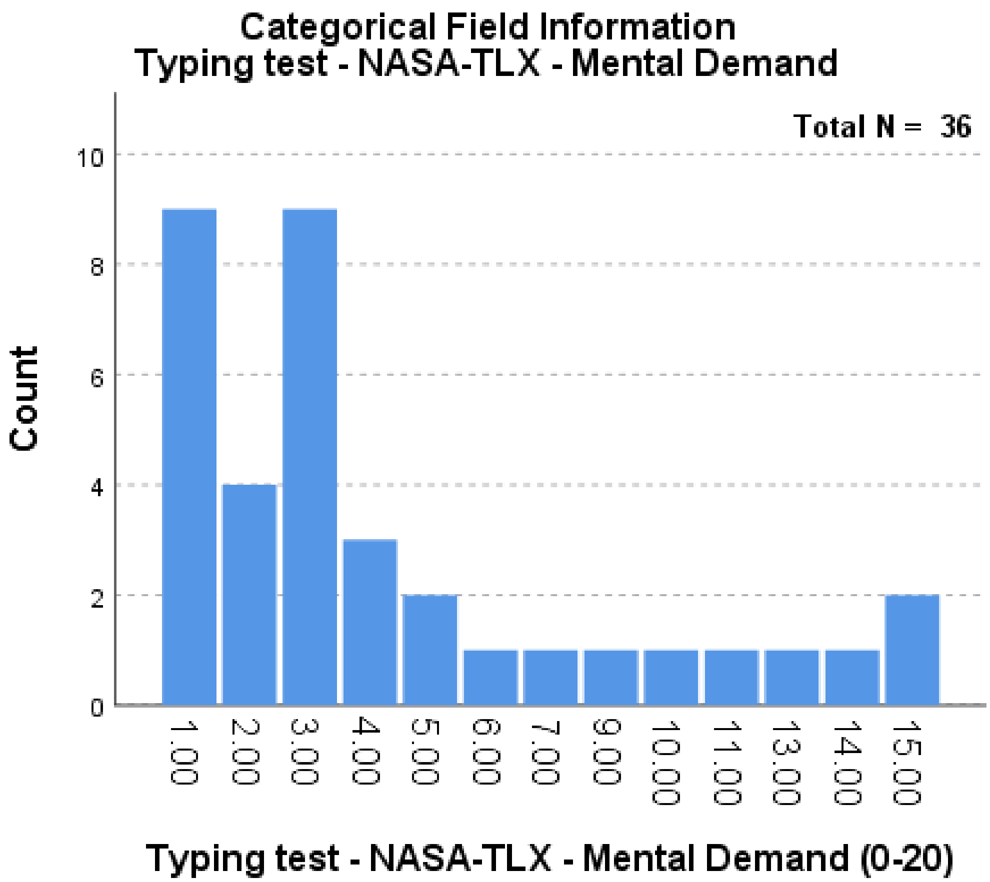
Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	6.001[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.306

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





Typing test - NASA-TLX - Mental Demand (0-20) field is ordinal but is treated as continuous in the test.

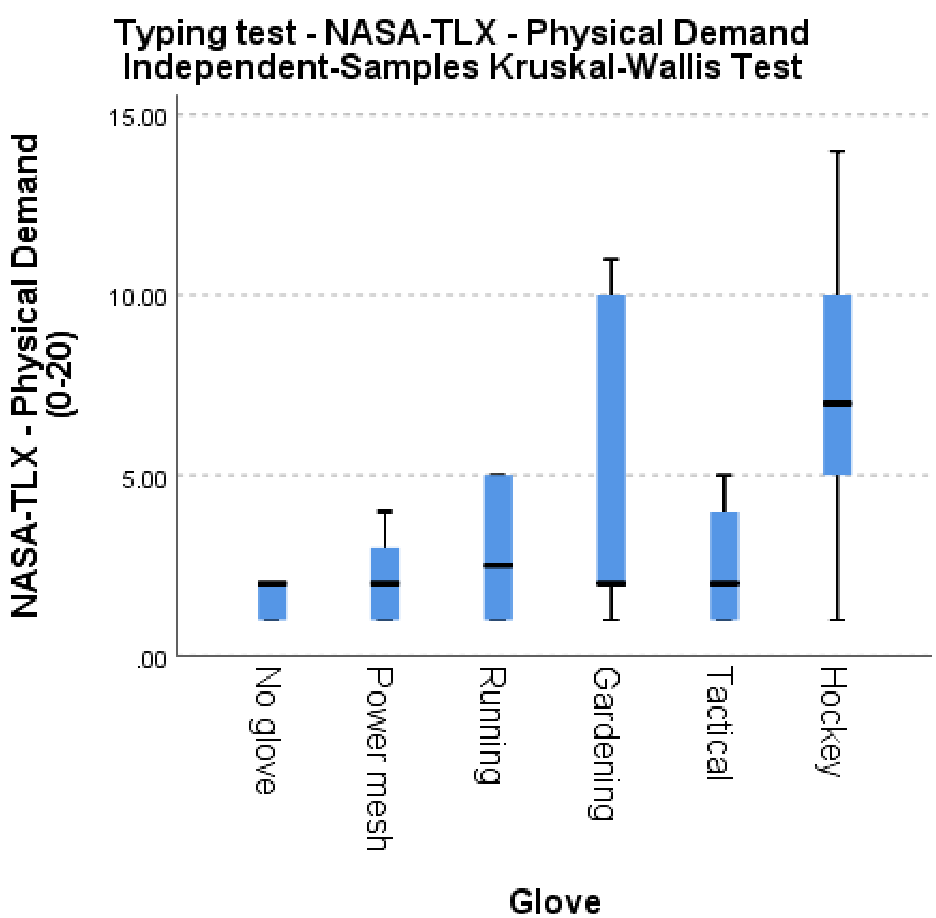
C.3.2 Typing test - NASA-TLX - Physical Demand

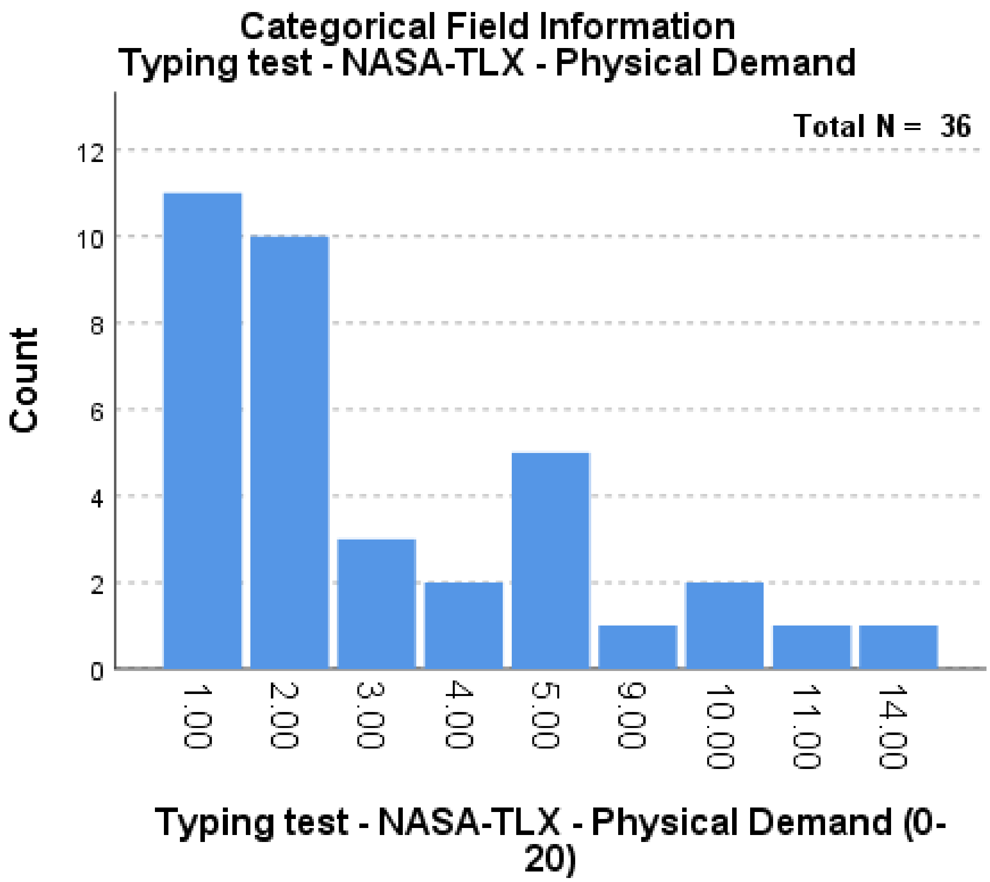
Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	7.573[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.181

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





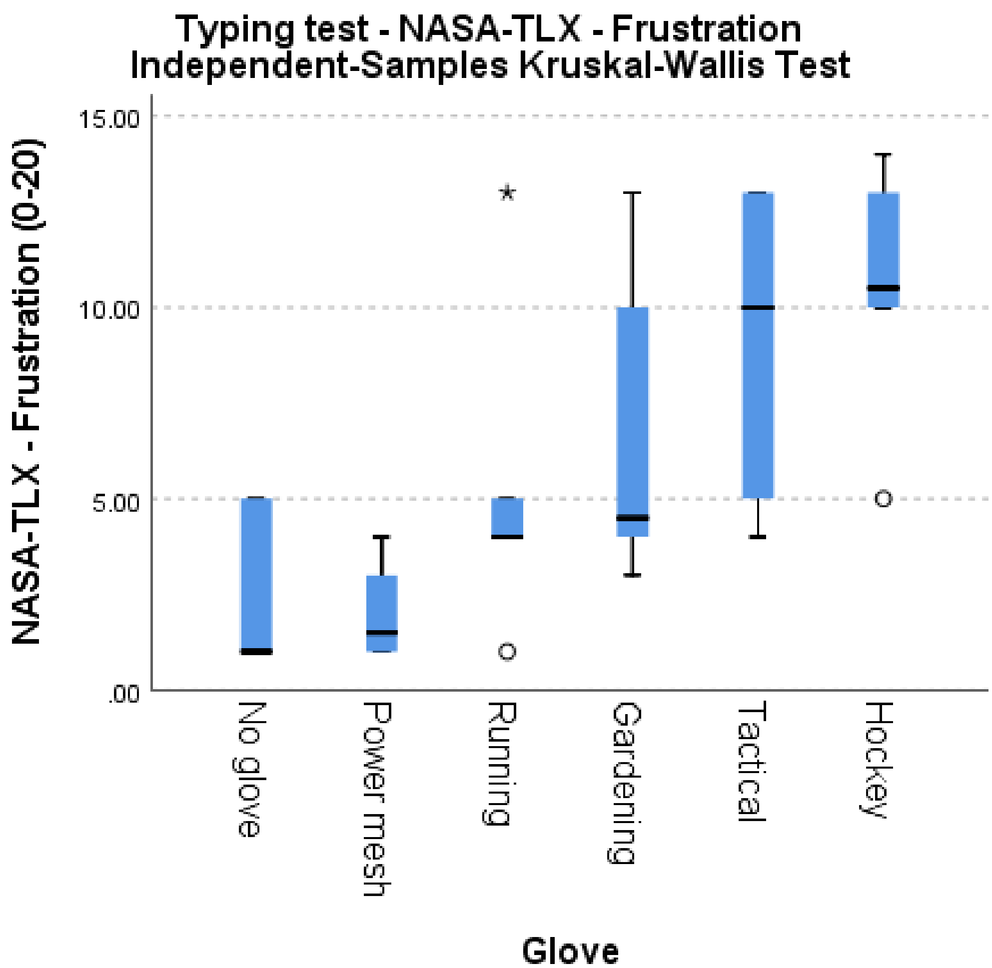
Typing test - NASA-TLX - Physical Demand (0-20) field is ordinal but is treated as continuous in the test.

C.3.3 Typing test - NASA-TLX - Frustration

Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	19.612[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.001

a. The test statistic is adjusted for ties.



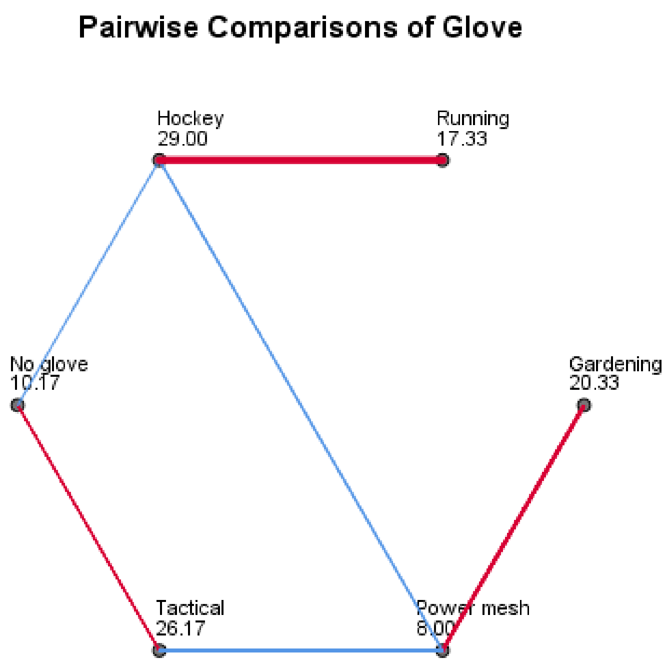
Pairwise Comparisons of Glove

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Power mesh-No glove	2.167	6.004	.361	.718	1.000
Power mesh-Running	-9.333	6.004	-1.555	.120	1.000
Power mesh-Gardening	-12.333	6.004	-2.054	.040	.599
Power mesh-Tactical	-18.167	6.004	-3.026	.002	.037
Power mesh-Hockey	-21.000	6.004	-3.498	.000	.007
No glove-Running	-7.167	6.004	-1.194	.233	1.000
No glove-Gardening	-10.167	6.004	-1.693	.090	1.000
No glove-Tactical	-16.000	6.004	-2.665	.008	.115
No glove-Hockey	-18.833	6.004	-3.137	.002	.026
Running-Gardening	-3.000	6.004	-.500	.617	1.000
Running-Tactical	-8.833	6.004	-1.471	.141	1.000
Running-Hockey	-11.667	6.004	-1.943	.052	.780
Gardening-Tactical	-5.833	6.004	-.972	.331	1.000
Gardening-Hockey	-8.667	6.004	-1.444	.149	1.000
Tactical-Hockey	-2.833	6.004	-.472	.637	1.000

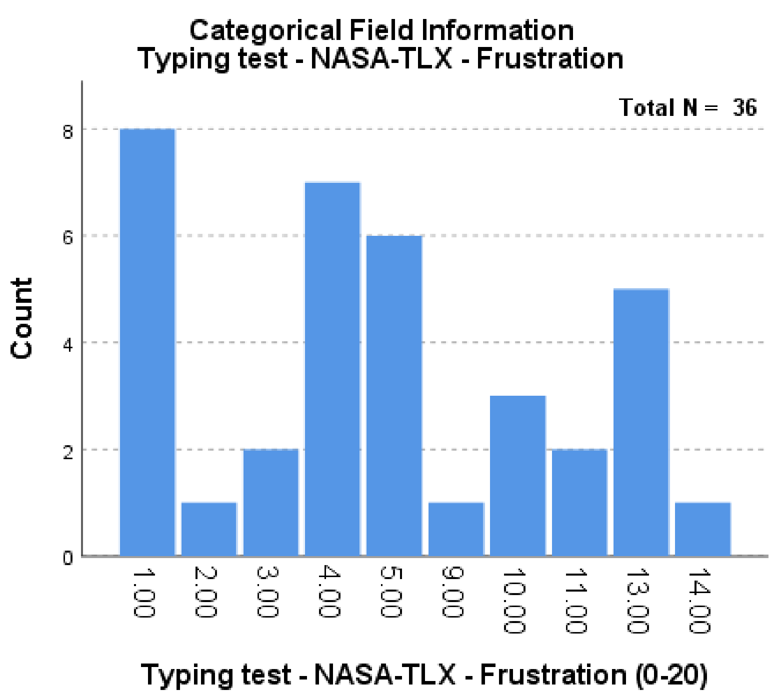
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



Each node shows the sample average rank of Glove.



Typing test - NASA-TLX - Frustration (0-20) field is ordinal but is treated as continuous in the test.

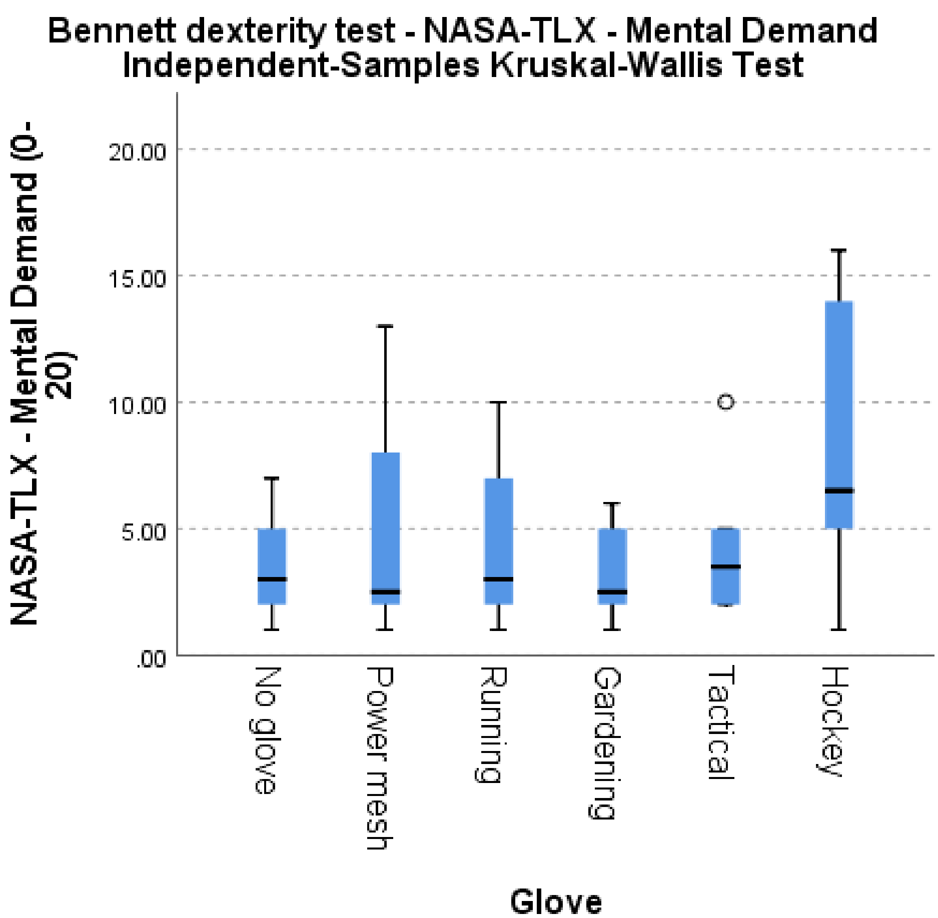
C.3.4 Bennett dexterity test - NASA-TLX - Mental Demand

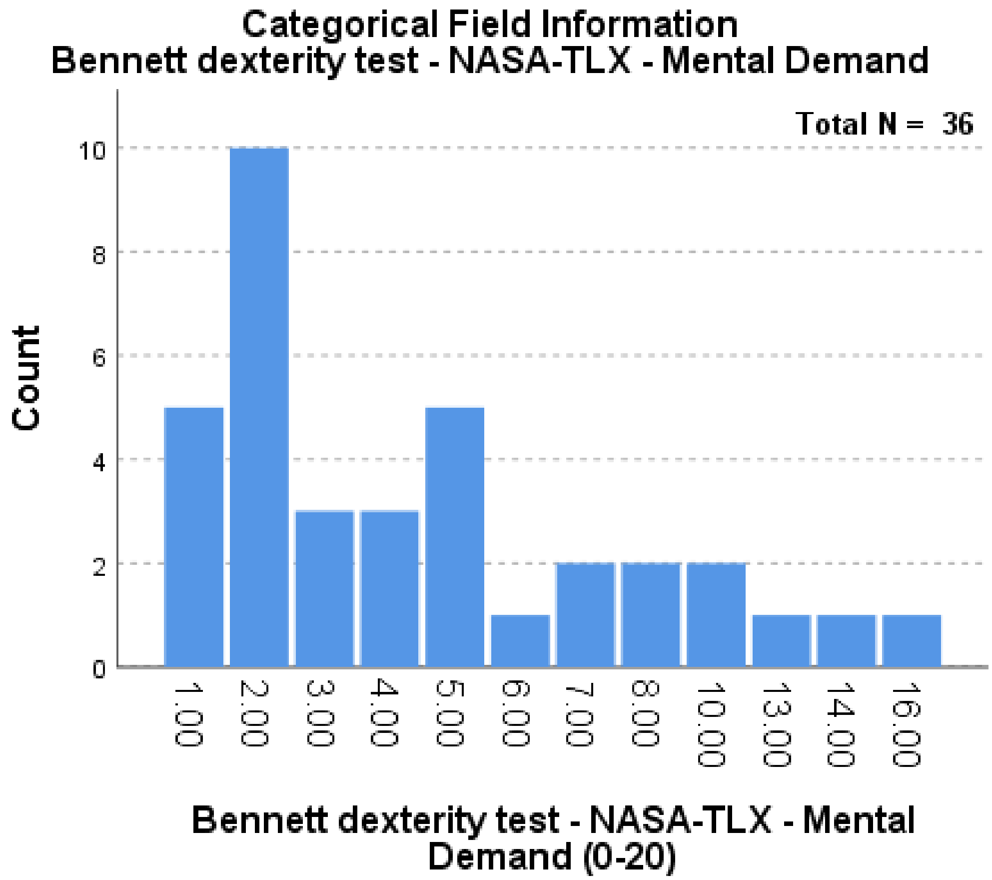
Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	3.663[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.599

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





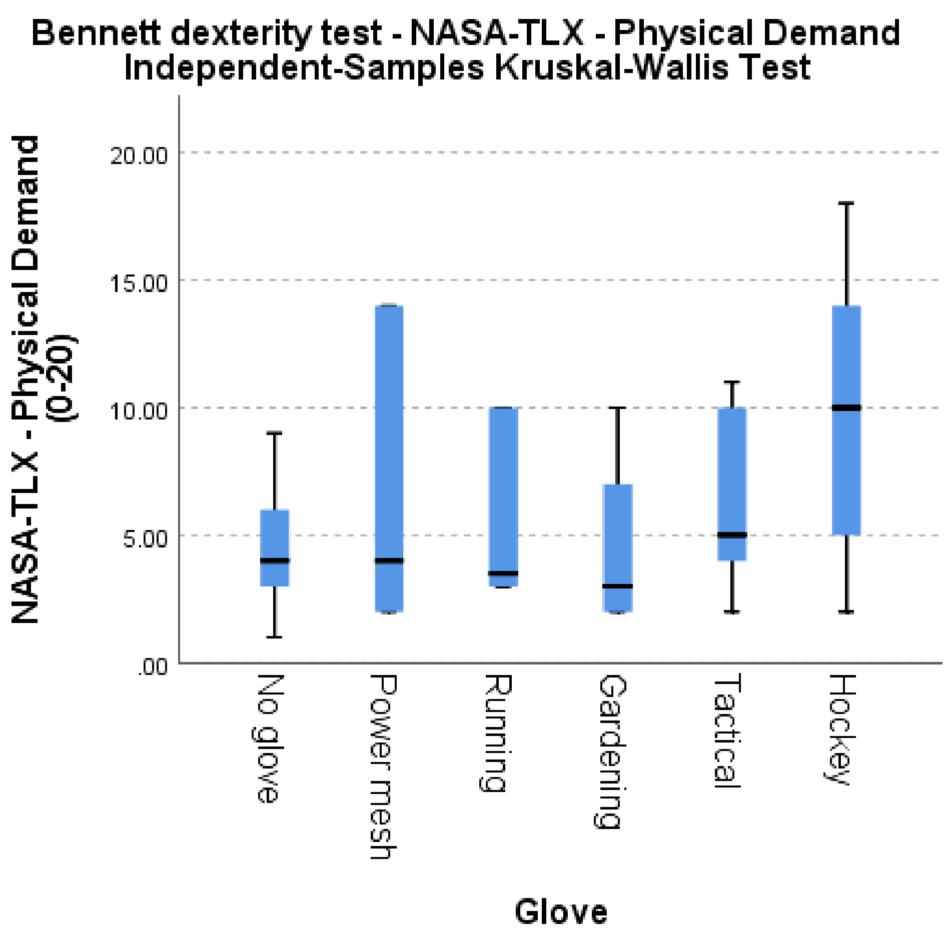
Bennett dexterity test - NASA-TLX - Mental Demand (0-20) field is ordinal but is treated as continuous in the test.

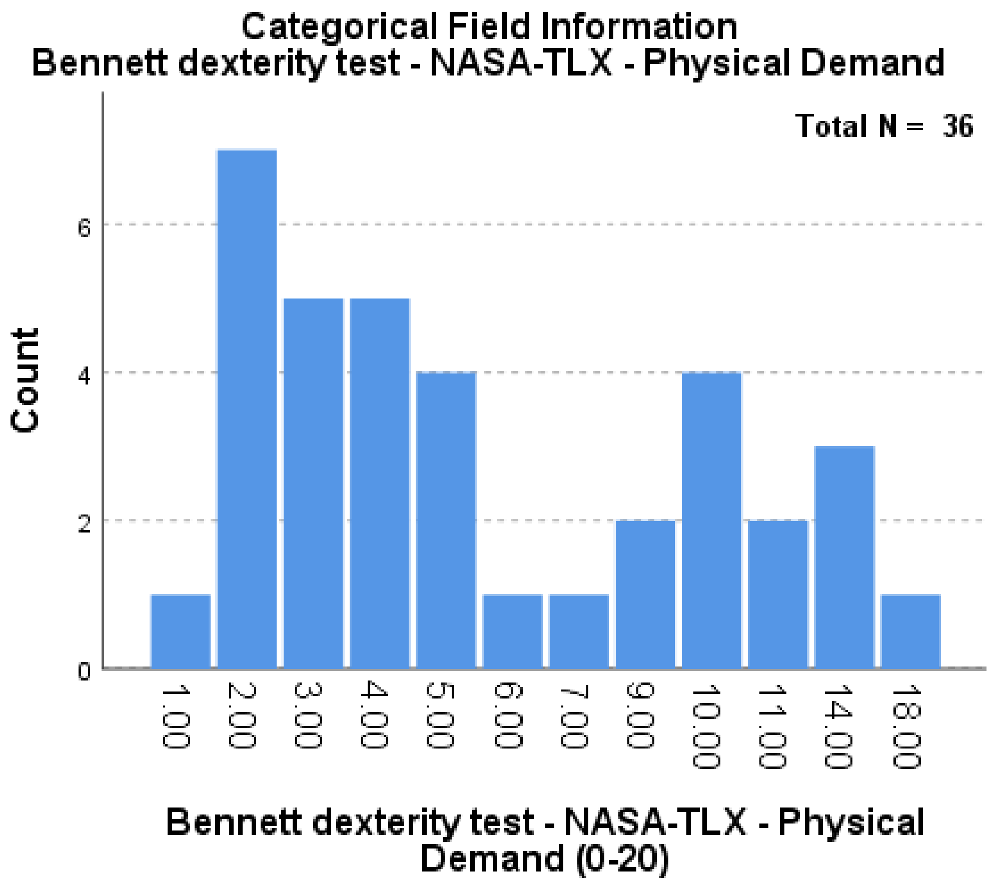
C.3.5 Bennett dexterity test - NASA-TLX - Physical Demand

Independent-Samples Kruskal-Wallis Test Summary	
Total N	36
Test Statistic	4.529[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.476

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





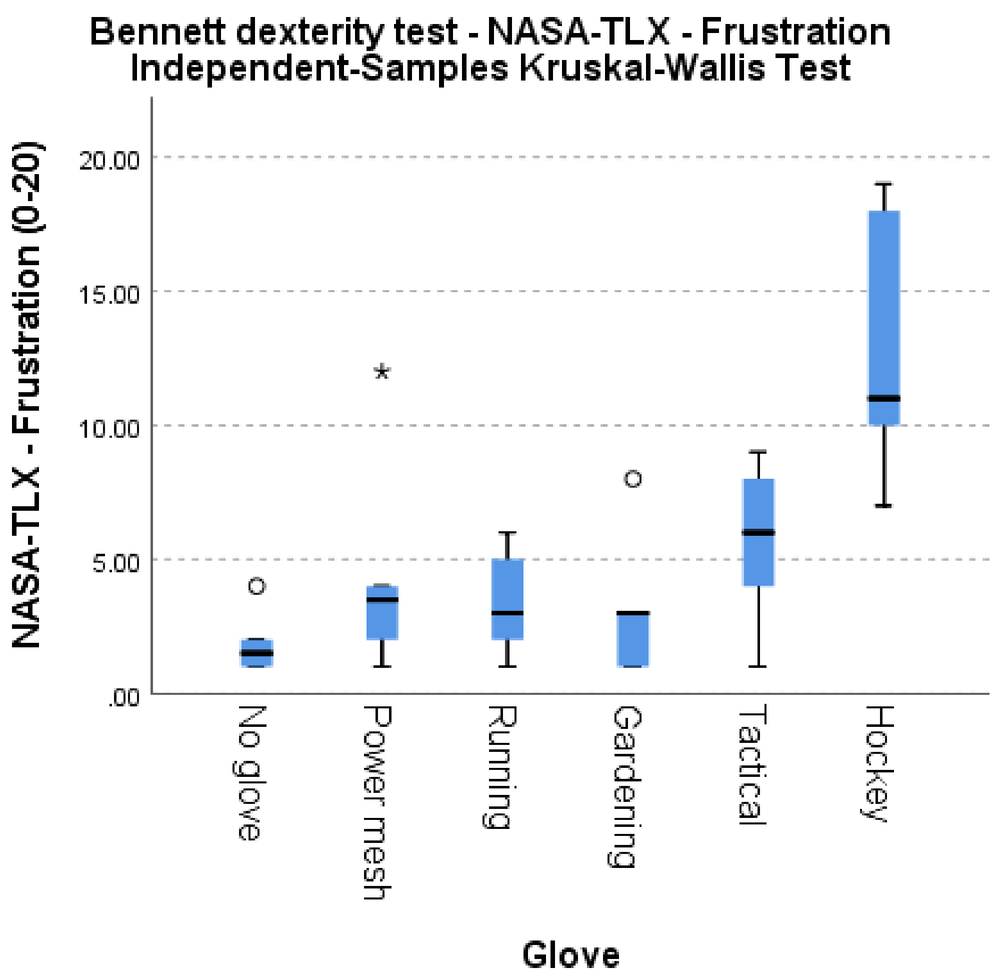
Bennett dexterity test - NASA-TLX - Physical Demand (0-20) field is ordinal but is treated as continuous in the test.

C.3.6 Bennett dexterity test - NASA-TLX - Frustration

Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	17.560[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.004

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

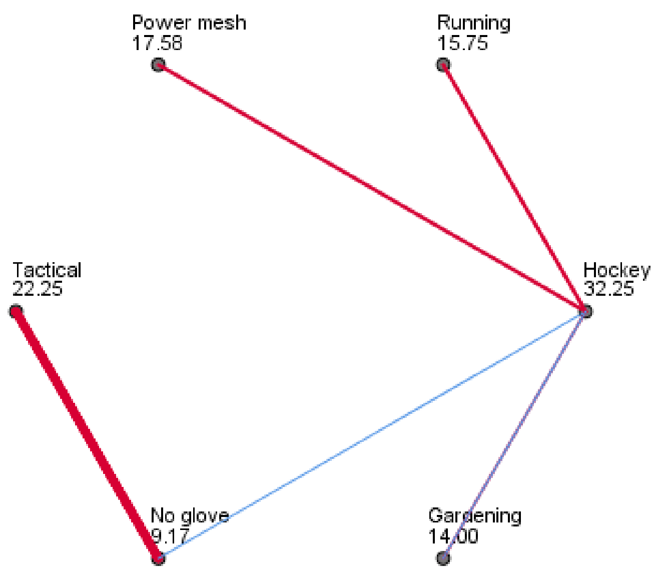
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
No glove-Gardening	-4.833	6.027	-.802	.423	1.000
No glove-Running	-6.583	6.027	-1.092	.275	1.000
No glove-Power mesh	-8.417	6.027	-1.397	.163	1.000
No glove-Tactical	-13.083	6.027	-2.171	.030	.449
No glove-Hockey	-23.083	6.027	-3.830	.000	.002
Gardening-Running	1.750	6.027	.290	.772	1.000
Gardening-Power mesh	3.583	6.027	.595	.552	1.000
Gardening-Tactical	-8.250	6.027	-1.369	.171	1.000
Gardening-Hockey	-18.250	6.027	-3.028	.002	.037
Running-Power mesh	1.833	6.027	.304	.761	1.000
Running-Tactical	-6.500	6.027	-1.079	.281	1.000
Running-Hockey	-16.500	6.027	-2.738	.006	.093
Power mesh-Tactical	-4.667	6.027	-.774	.439	1.000
Power mesh-Hockey	-14.667	6.027	-2.434	.015	.224
Tactical-Hockey	-10.000	6.027	-1.659	.097	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

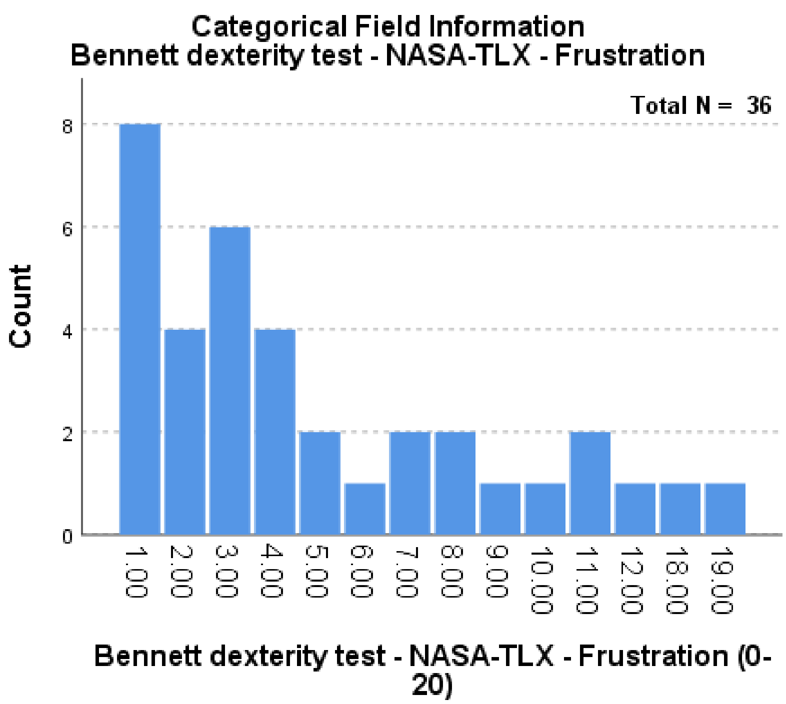
Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.



Bennett dexterity test - NASA-TLX - Frustration (0-20) field is ordinal but is treated as continuous in the test.

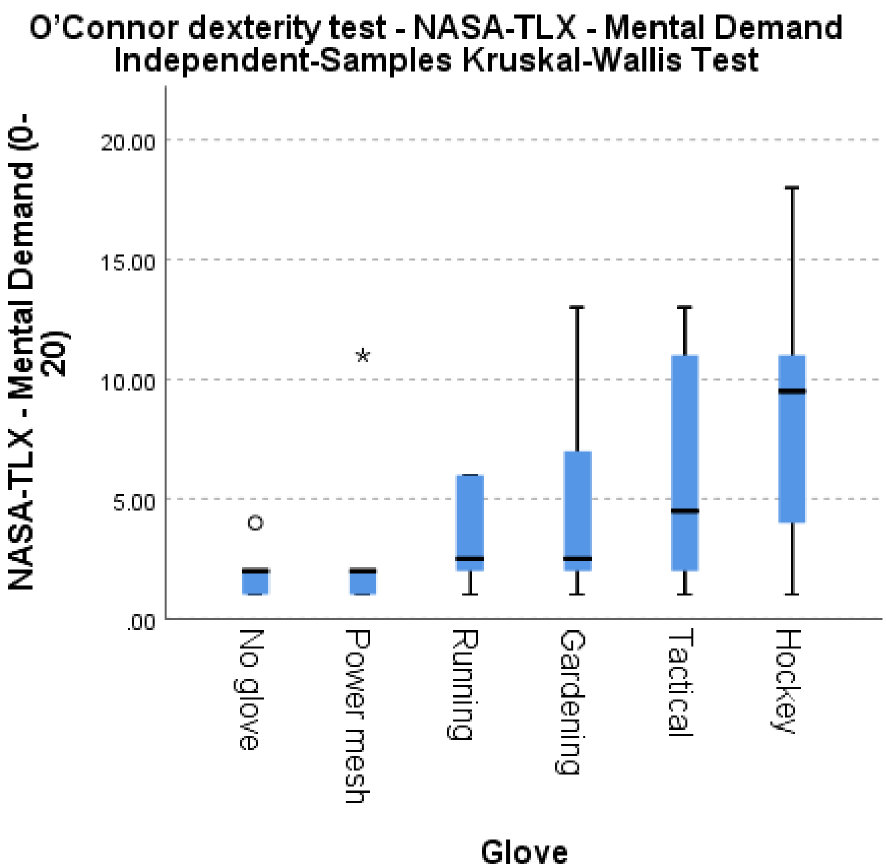
C.3.7 O'Connor dexterity test - NASA-TLX - Mental Demand

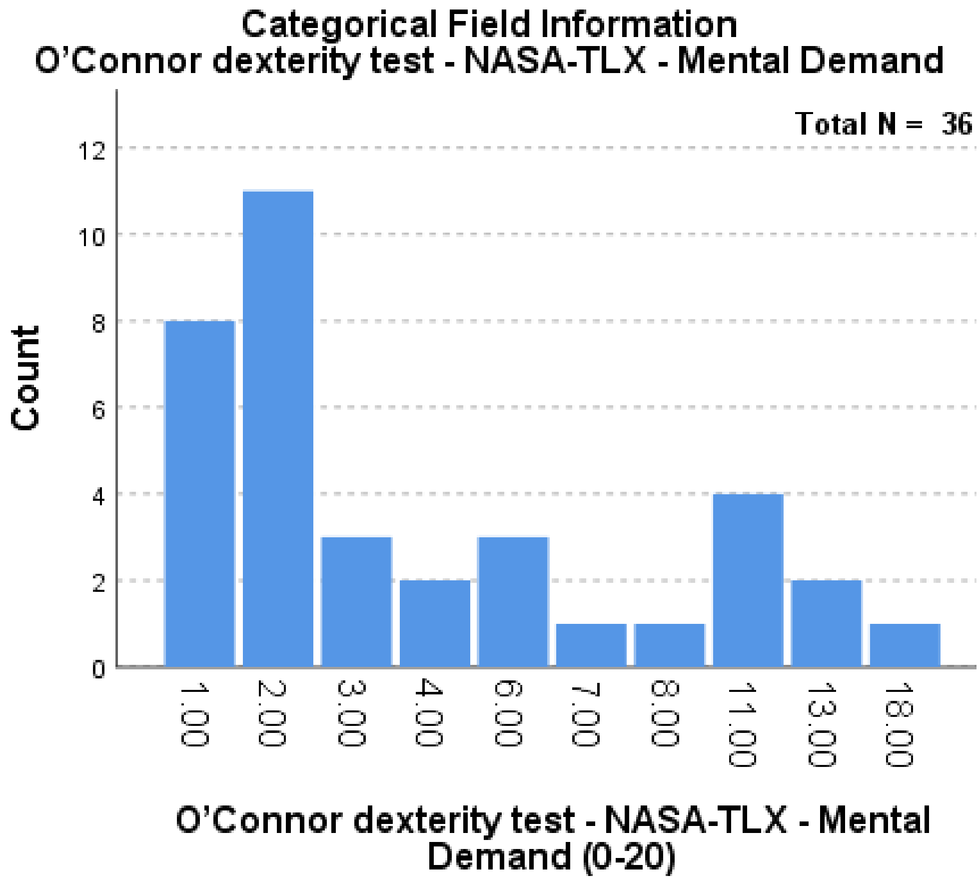
Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	7.278[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.201

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





O'Connor dexterity test - NASA-TLX - Mental Demand (0-20) field is ordinal but is treated as continuous in the test.

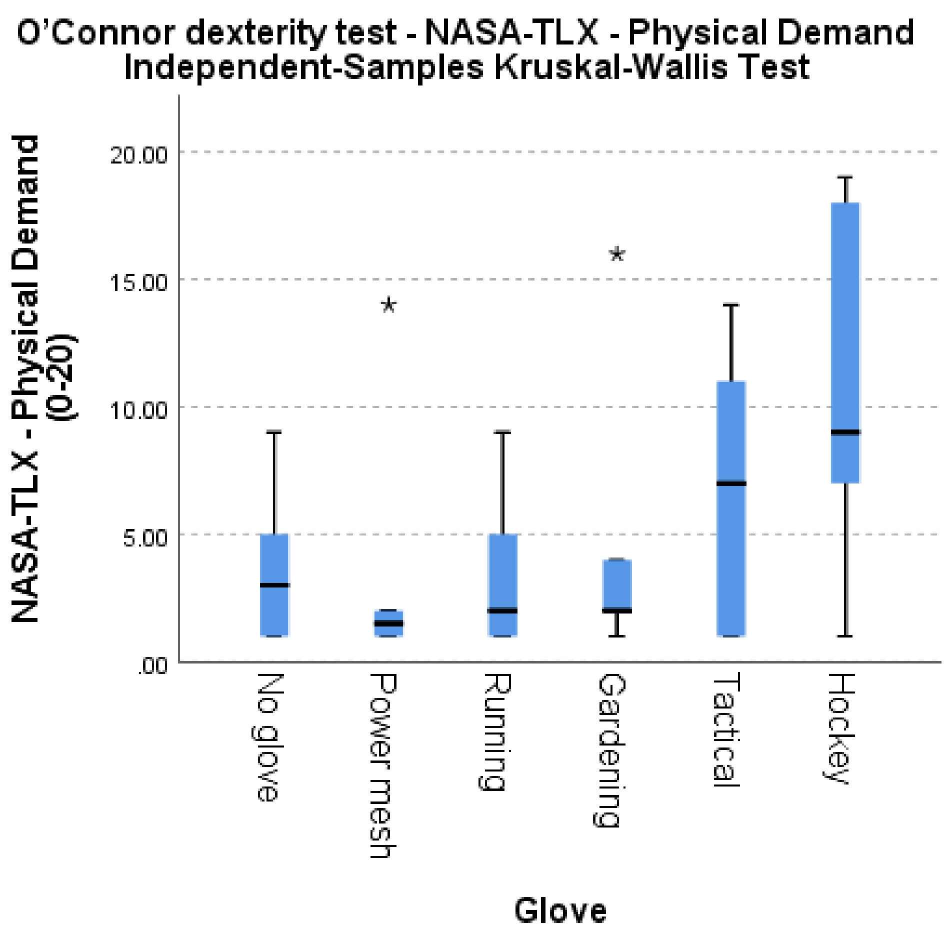
C.3.8 O'Connor dexterity test - NASA-TLX - Physical Demand

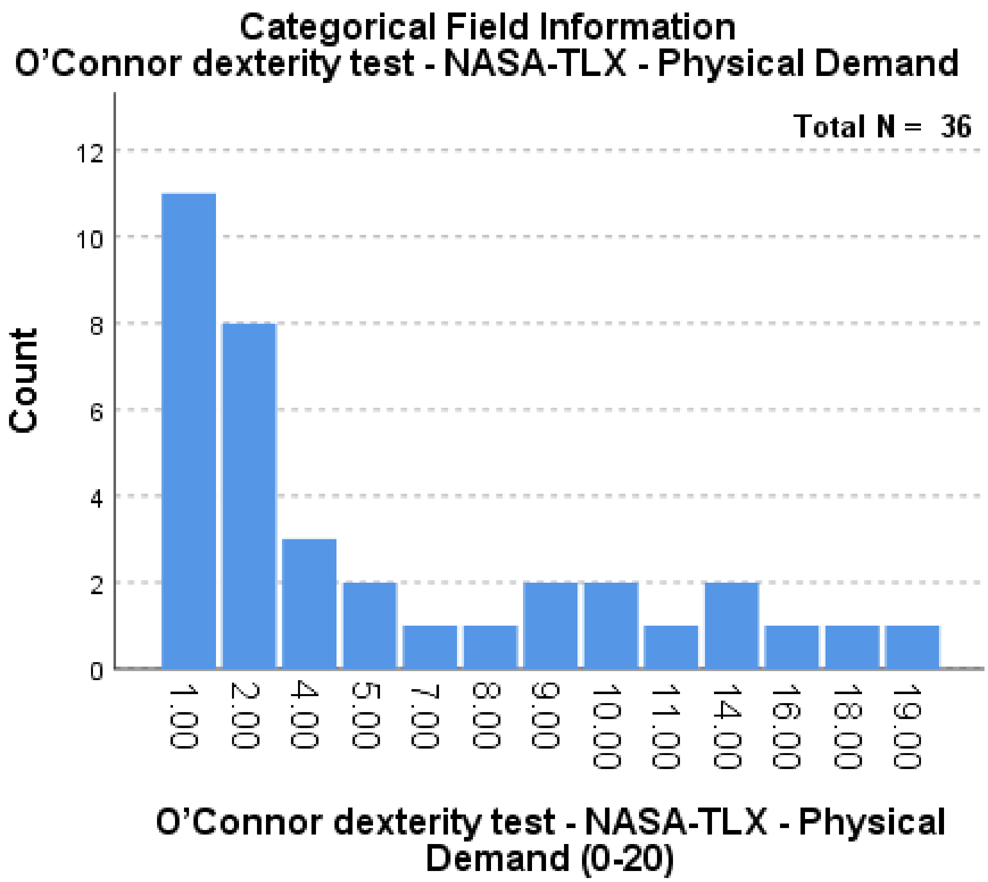
Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	5.774[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.329

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





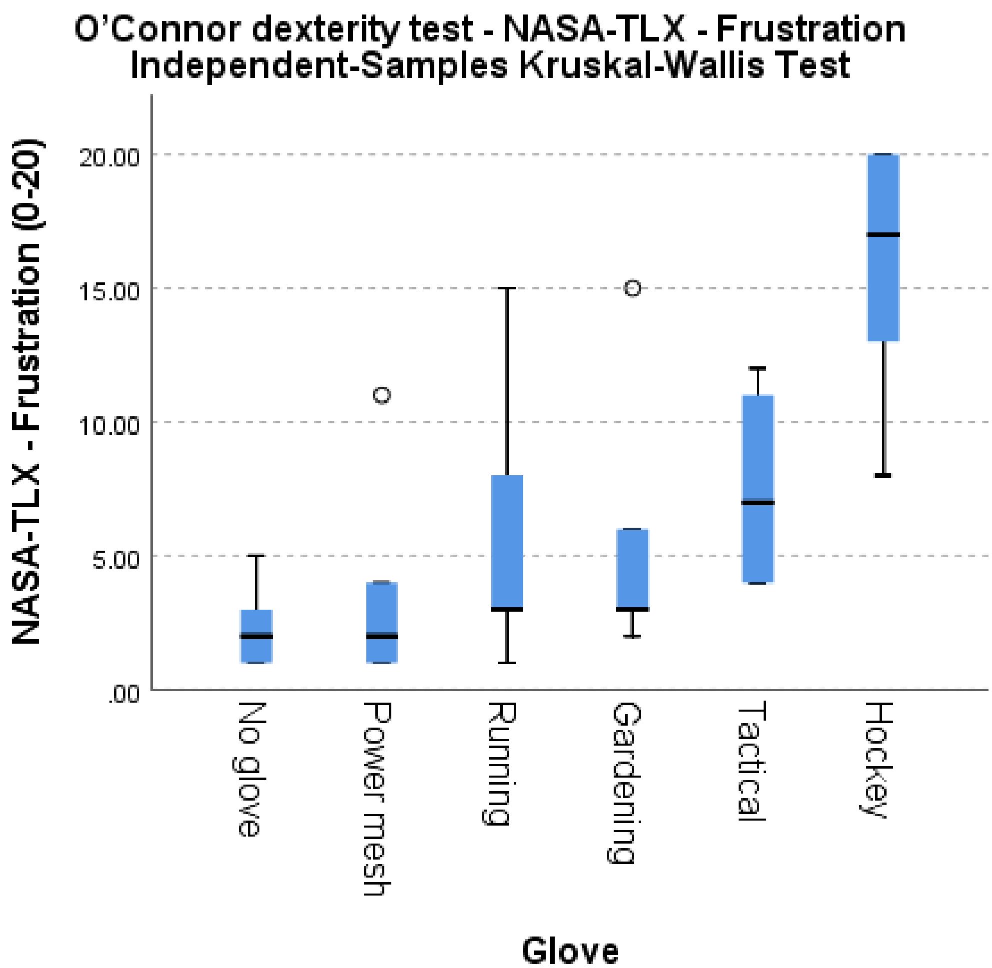
O'Connor dexterity test - NASA-TLX - Physical Demand (0-20) field is ordinal but is treated as continuous in the test.

C.3.9 O'Connor dexterity test - NASA-TLX - Frustration (0-20) across Glove

Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	18.138[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.003

a. The test statistic is adjusted for ties.



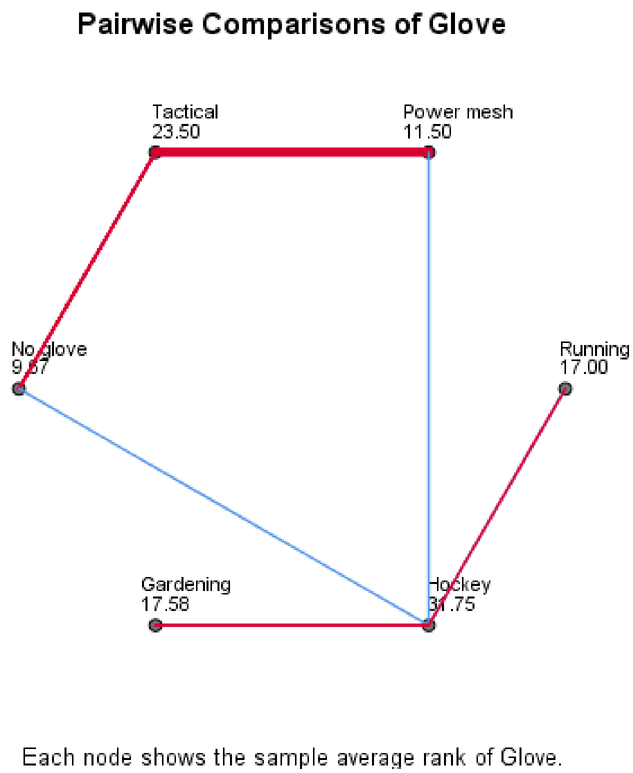
Pairwise Comparisons of Glove

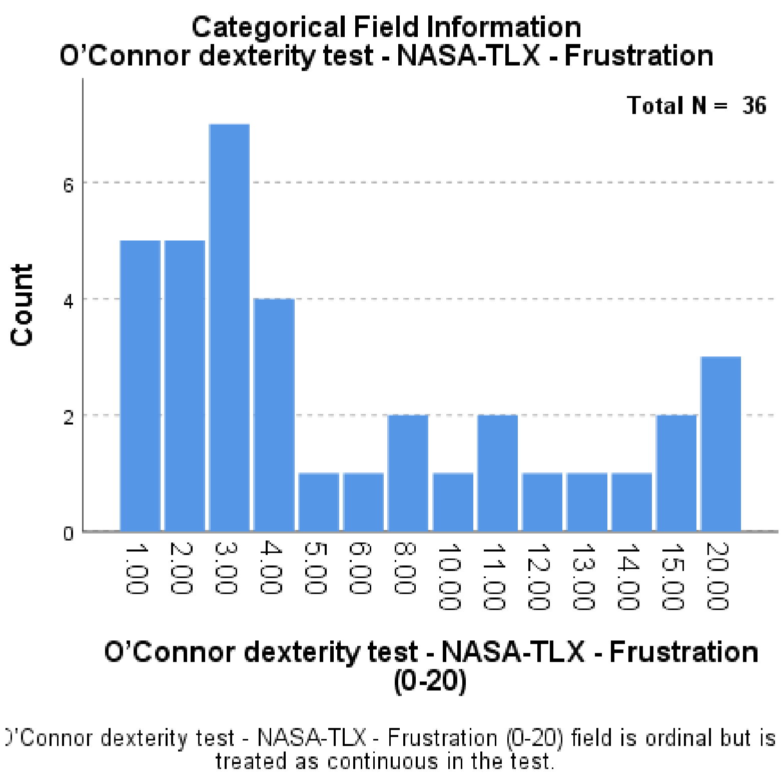
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
No glove-Power mesh	-1.833	6.038	-.304	.761	1.000
No glove-Running	-7.333	6.038	-1.214	.225	1.000
No glove-Gardening	-7.917	6.038	-1.311	.190	1.000
No glove-Tactical	-13.833	6.038	-2.291	.022	.330
No glove-Hockey	-22.083	6.038	-3.657	.000	.004
Power mesh-Running	-5.500	6.038	-.911	.362	1.000
Power mesh-Gardening	-6.083	6.038	-1.007	.314	1.000
Power mesh-Tactical	-12.000	6.038	-1.987	.047	.703
Power mesh-Hockey	-20.250	6.038	-3.354	.001	.012
Running-Gardening	-.583	6.038	-.097	.923	1.000
Running-Tactical	-6.500	6.038	-1.076	.282	1.000
Running-Hockey	-14.750	6.038	-2.443	.015	.219
Gardening-Tactical	-5.917	6.038	-.980	.327	1.000
Gardening-Hockey	-14.167	6.038	-2.346	.019	.285
Tactical-Hockey	-8.250	6.038	-1.366	.172	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.


Pairwise Comparisons of Glove
Tactical
23.50
Power mesh
11.50
No glove
9.67
Running
17.00
Gardening
17.58
Hockey
31.75
Each node shows the sample average rank of Glove.

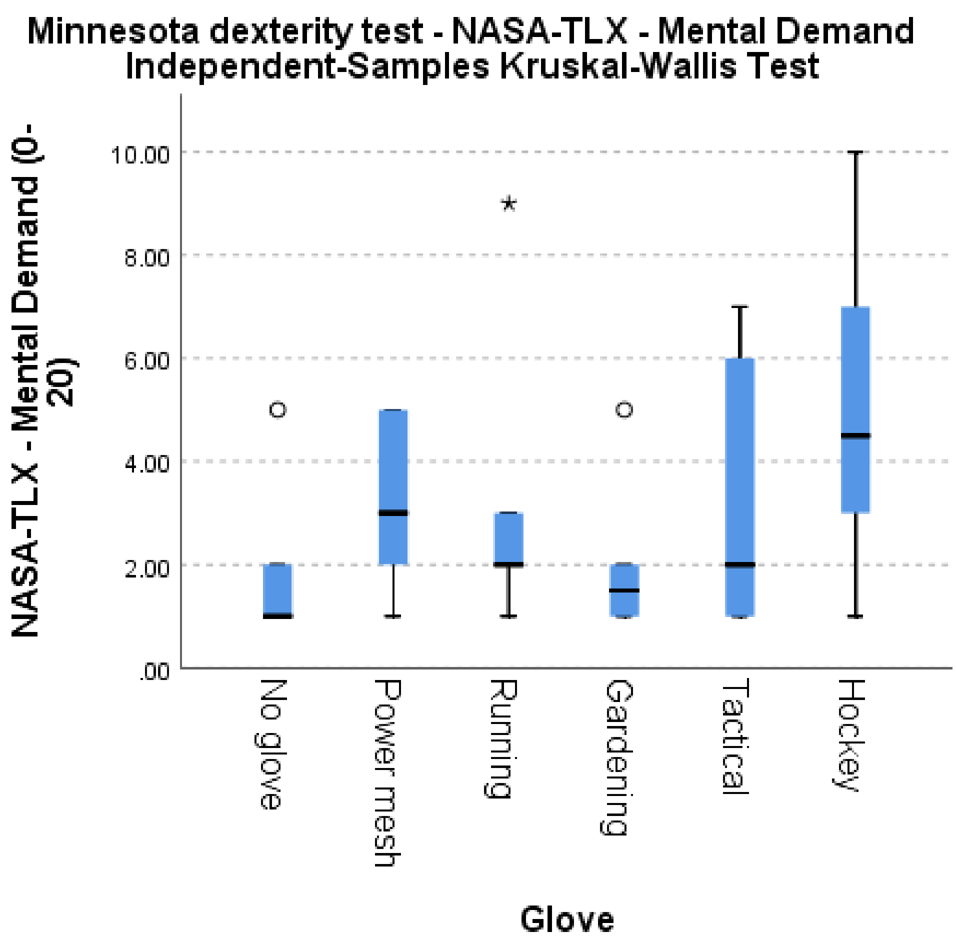

Categorical Field Information
O'Connor dexterity test - NASA-TLX - Frustration
Total N = 36
Count
0
2
4
6
1.00
2.00
3.00
4.00
5.00
6.00
8.00
10.00
11.00
12.00
13.00
14.00
15.00
20.00
O'Connor dexterity test - NASA-TLX - Frustration (0-20)
O'Connor dexterity test - NASA-TLX - Frustration (0-20) field is ordinal but is treated as continuous in the test.

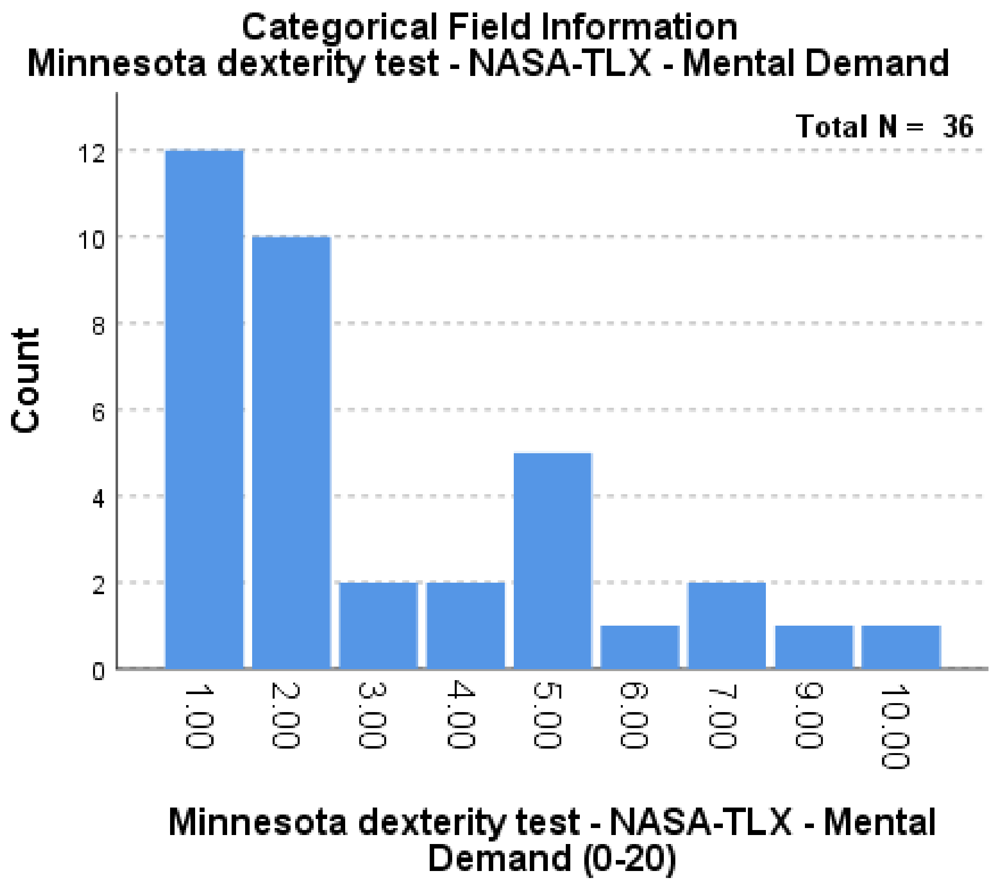
C.3.10 Minnesota dexterity test - NASA-TLX - Mental Demand

Independent-Samples Kruskal-Wallis Test Summary	
Total N	36
Test Statistic	6.927[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.226

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





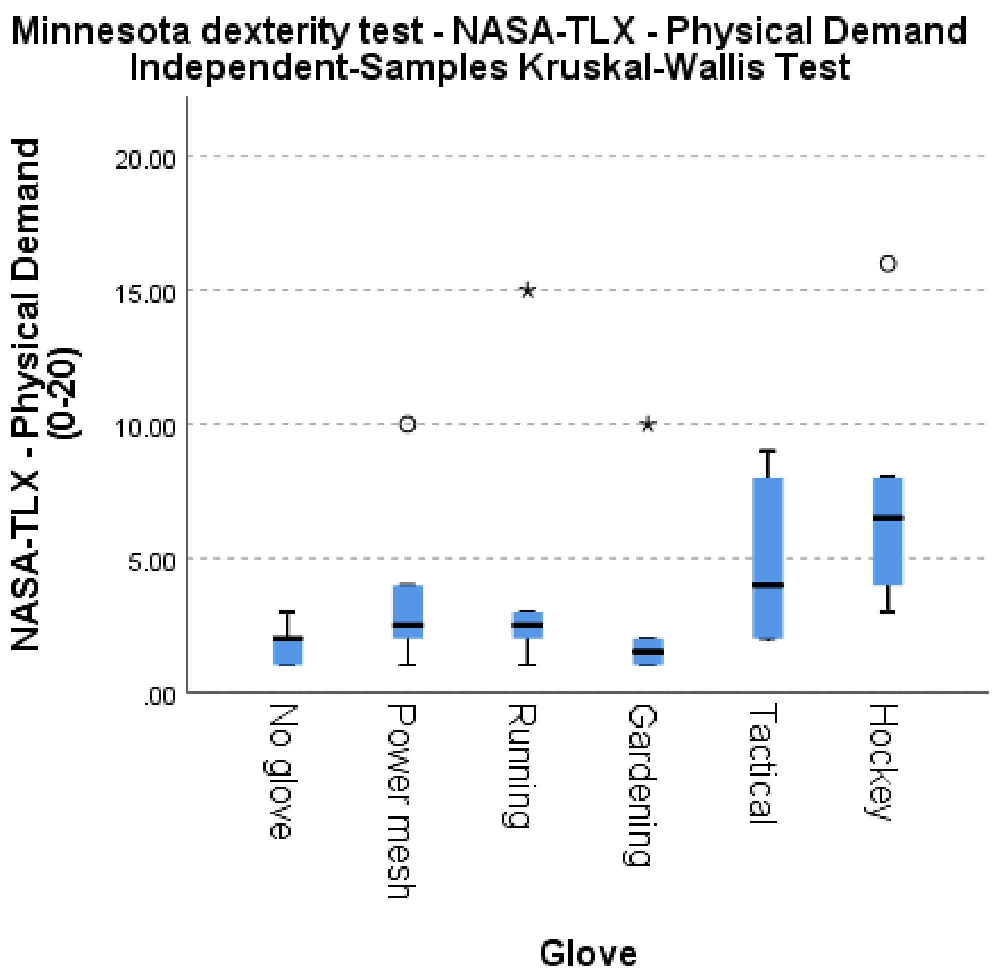
Minnesota dexterity test - NASA-TLX - Mental Demand (0-20) field is ordinal but is treated as continuous in the test.

C.3.11 Minnesota dexterity test - NASA-TLX - Physical Demand

Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	11.886[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.036

a. The test statistic is adjusted for ties.



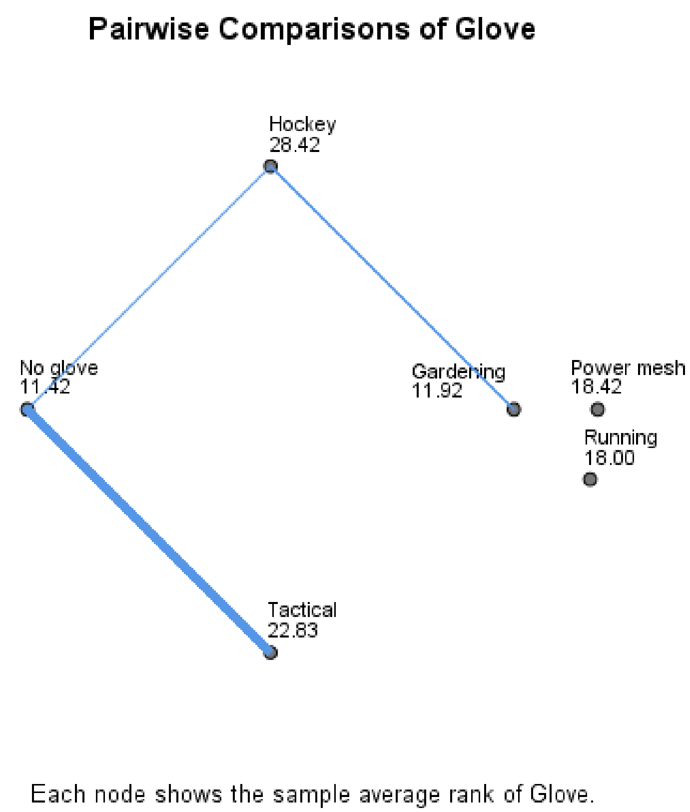
Pairwise Comparisons of Glove

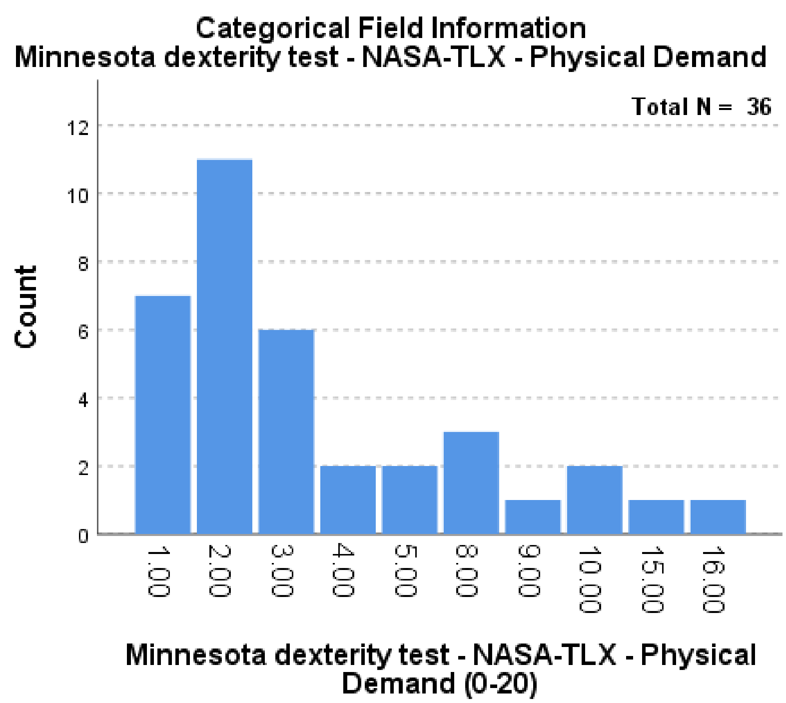
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
No glove-Gardening	-.500	5.957	-.084	.933	1.000
No glove-Running	-6.583	5.957	-1.105	.269	1.000
No glove-Power mesh	-7.000	5.957	-1.175	.240	1.000
No glove-Tactical	-11.417	5.957	-1.917	.055	.829
No glove-Hockey	-17.000	5.957	-2.854	.004	.065
Gardening-Running	6.083	5.957	1.021	.307	1.000
Gardening-Power mesh	6.500	5.957	1.091	.275	1.000
Gardening-Tactical	-10.917	5.957	-1.833	.067	1.000
Gardening-Hockey	-16.500	5.957	-2.770	.006	.084
Running-Power mesh	.417	5.957	.070	.944	1.000
Running-Tactical	-4.833	5.957	-.811	.417	1.000
Running-Hockey	-10.417	5.957	-1.749	.080	1.000
Power mesh-Tactical	-4.417	5.957	-.741	.458	1.000
Power mesh-Hockey	-10.000	5.957	-1.679	.093	1.000
Tactical-Hockey	-5.583	5.957	-.937	.349	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



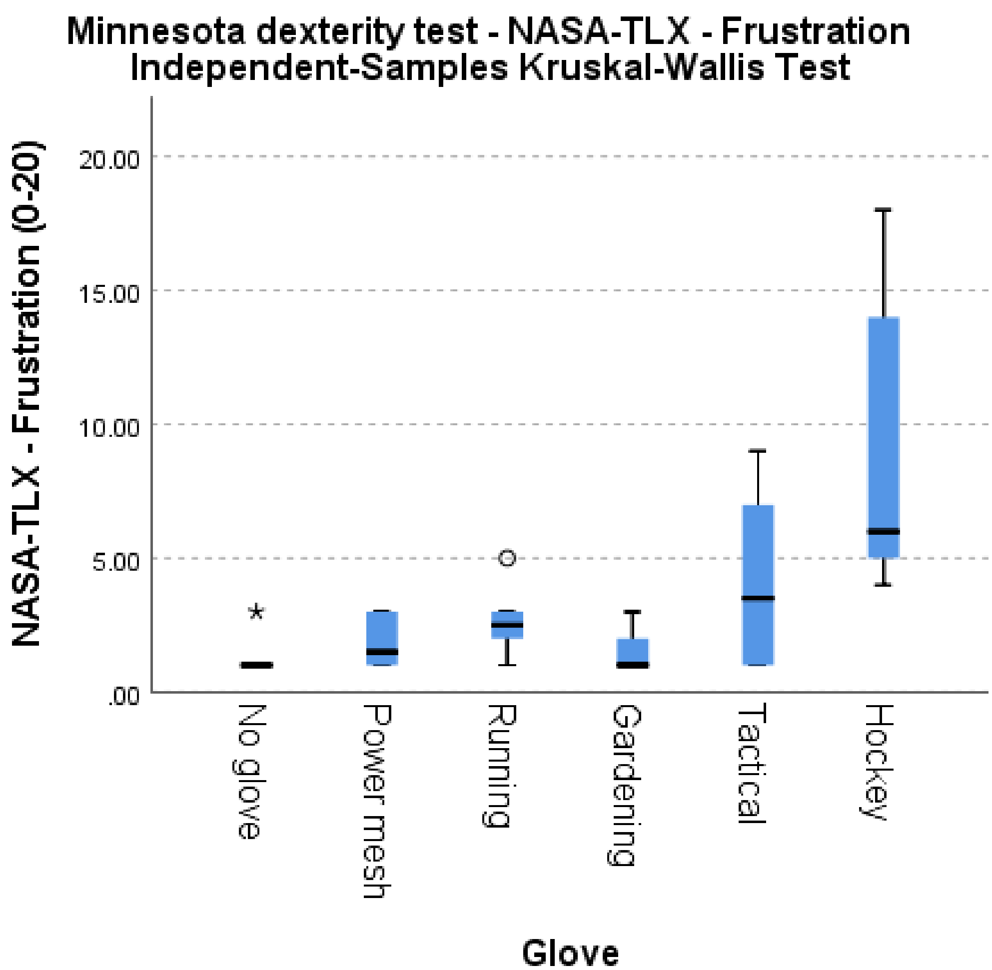


Minnesota dexterity test - NASA-TLX - Physical Demand (0-20) field is ordinal but is treated as continuous in the test.

C.3.12 Minnesota dexterity test - NASA-TLX - Frustration

Independent-Samples Kruskal-Wallis Test Summary	
Total N	36
Test Statistic	18.447[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.002

a. The test statistic is adjusted for ties.



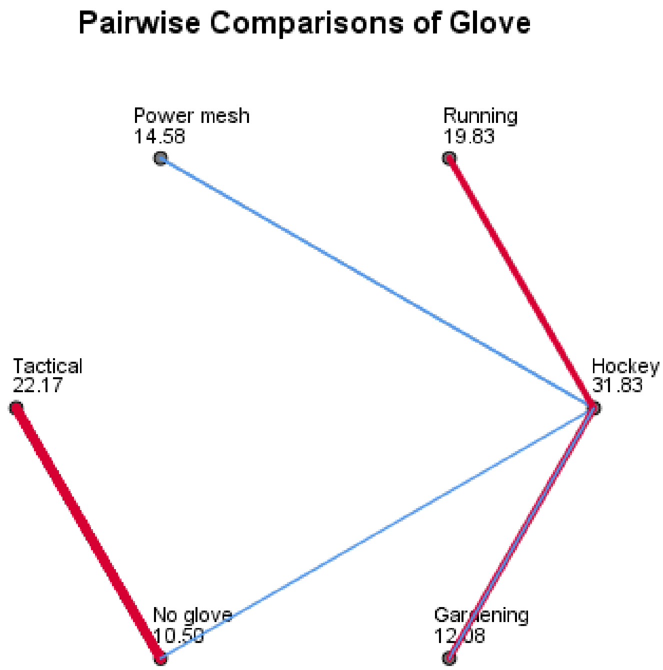
Pairwise Comparisons of Glove

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
No glove-Gardening	-1.583	5.830	-.272	.786	1.000
No glove-Power mesh	-4.083	5.830	-.700	.484	1.000
No glove-Running	-9.333	5.830	-1.601	.109	1.000
No glove-Tactical	-11.667	5.830	-2.001	.045	.681
No glove-Hockey	-21.333	5.830	-3.659	.000	.004
Gardening-Power mesh	2.500	5.830	.429	.668	1.000
Gardening-Running	7.750	5.830	1.329	.184	1.000
Gardening-Tactical	-10.083	5.830	-1.730	.084	1.000
Gardening-Hockey	-19.750	5.830	-3.388	.001	.011
Power mesh-Running	-5.250	5.830	-.900	.368	1.000
Power mesh-Tactical	-7.583	5.830	-1.301	.193	1.000
Power mesh-Hockey	-17.250	5.830	-2.959	.003	.046
Running-Tactical	-2.333	5.830	-.400	.689	1.000
Running-Hockey	-12.000	5.830	-2.058	.040	.593
Tactical-Hockey	-9.667	5.830	-1.658	.097	1.000

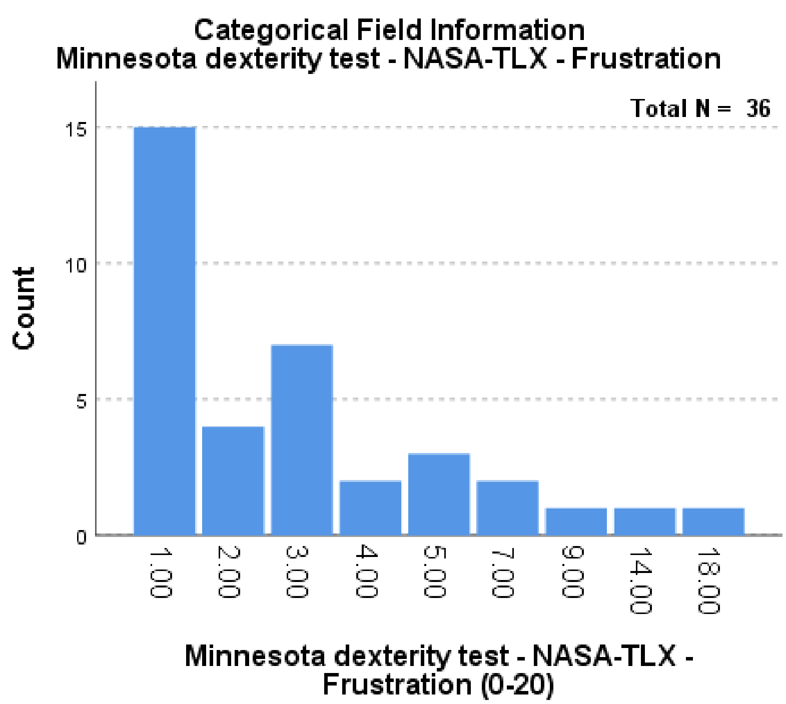
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



Each node shows the sample average rank of Glove.



Minnesota dexterity test - NASA-TLX - Frustration (0-20) field is ordinal but is treated as continuous in the test.

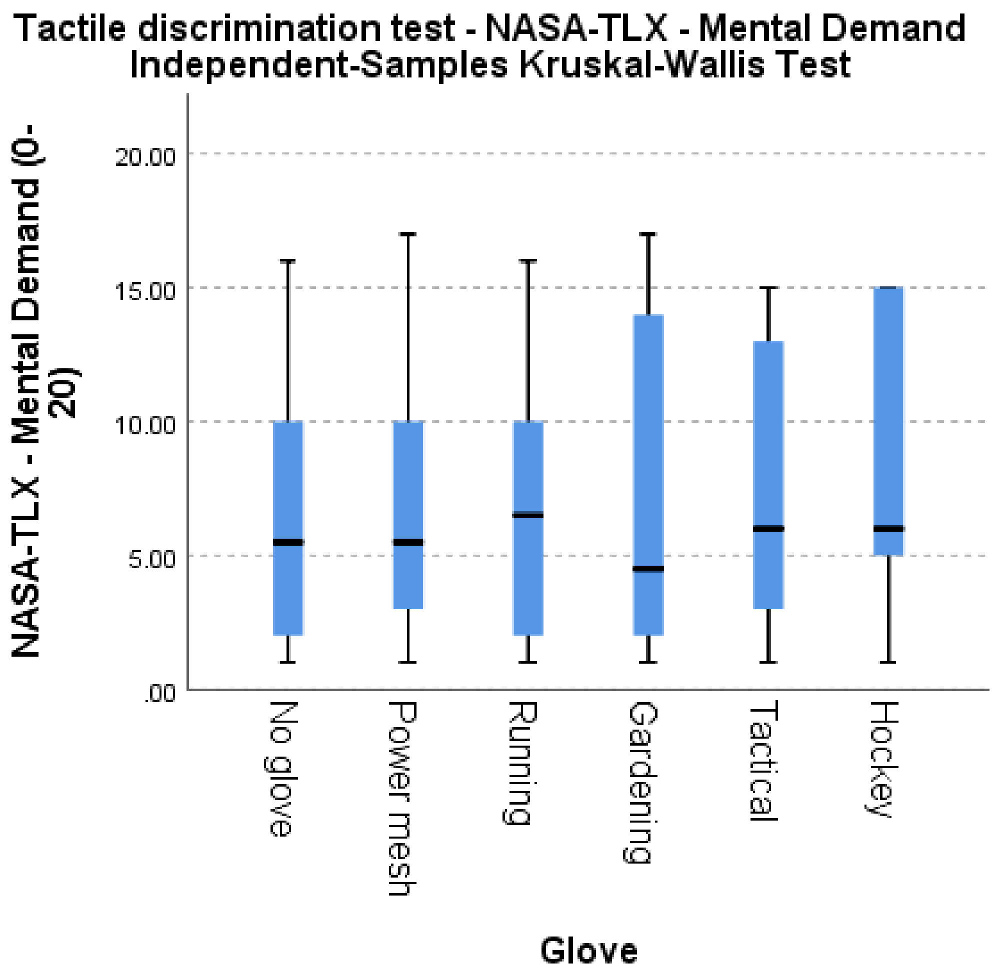
C.3.13 Tactile discrimination test - NASA-TLX - Mental Demand

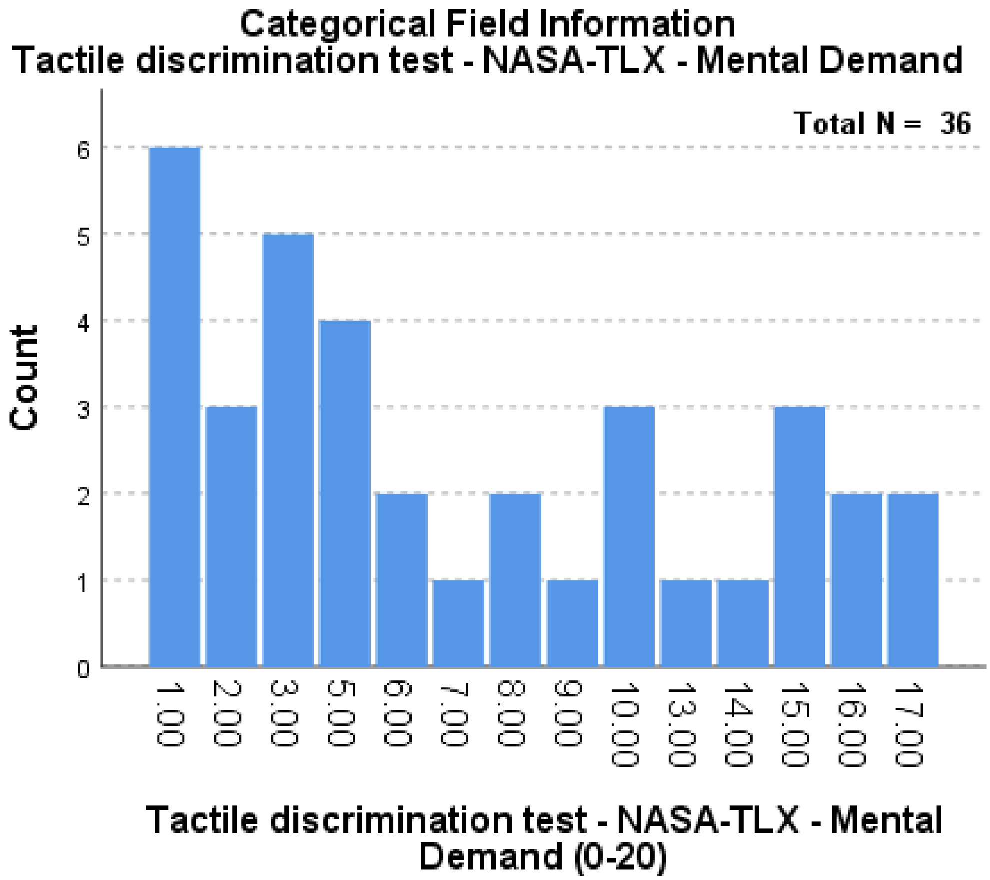
Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	.182[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.999

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.


Tactile discrimination test - NASA-TLX - Mental Demand
Independent-Samples Kruskal-Wallis Test
NASA-TLX - Mental Demand (0-20)
20.00
15.00
10.00
5.00
.00
No glove
Power mesh
Running
Gardening
Tactical
Hockey
Glove



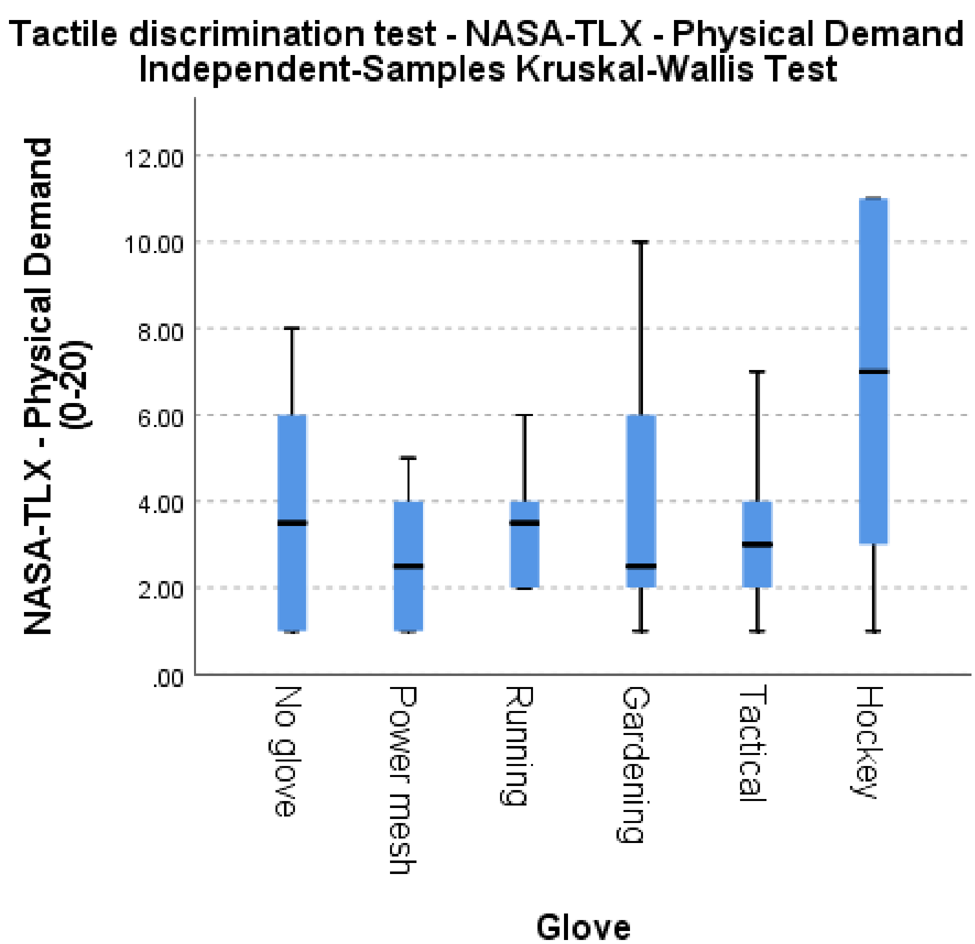
Tactile discrimination test - NASA-TLX - Mental Demand (0-20) field is ordinal but is treated as continuous in the test.

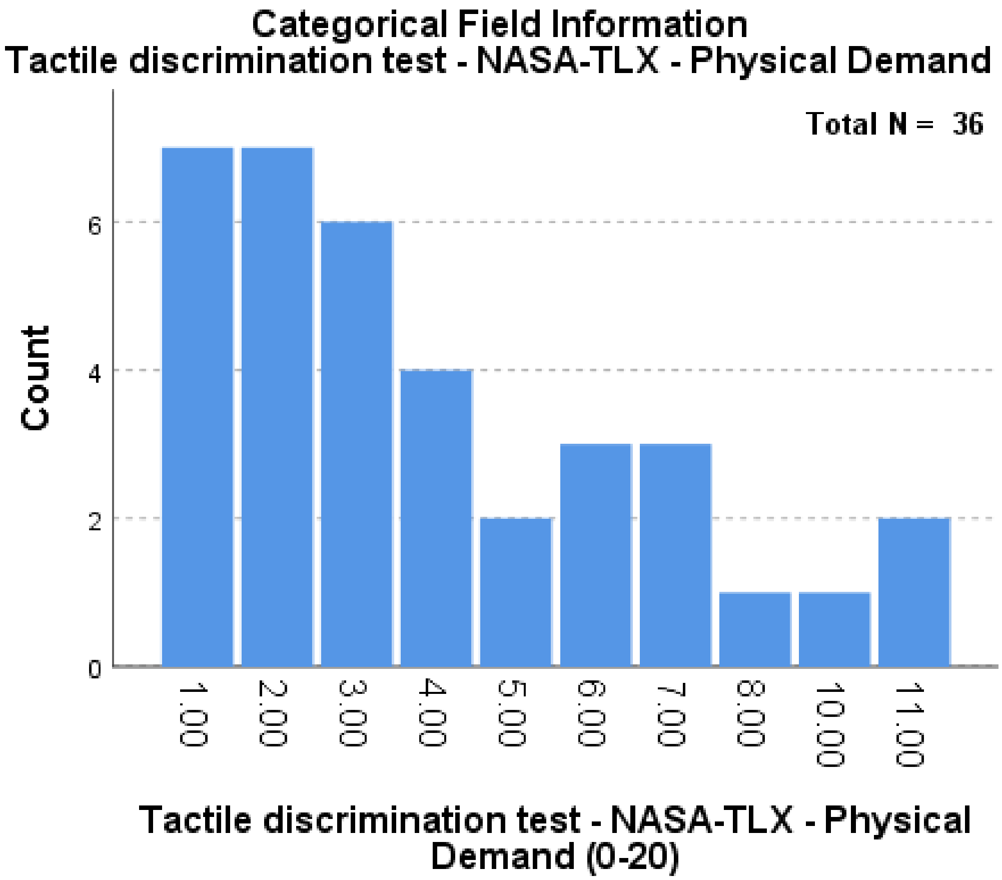
C.3.14 Tactile discrimination test - NASA-TLX - Physical Demand

Independent-Samples Kruskal-Wallis Test Summary	
Total N	36
Test Statistic	4.164[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.526

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





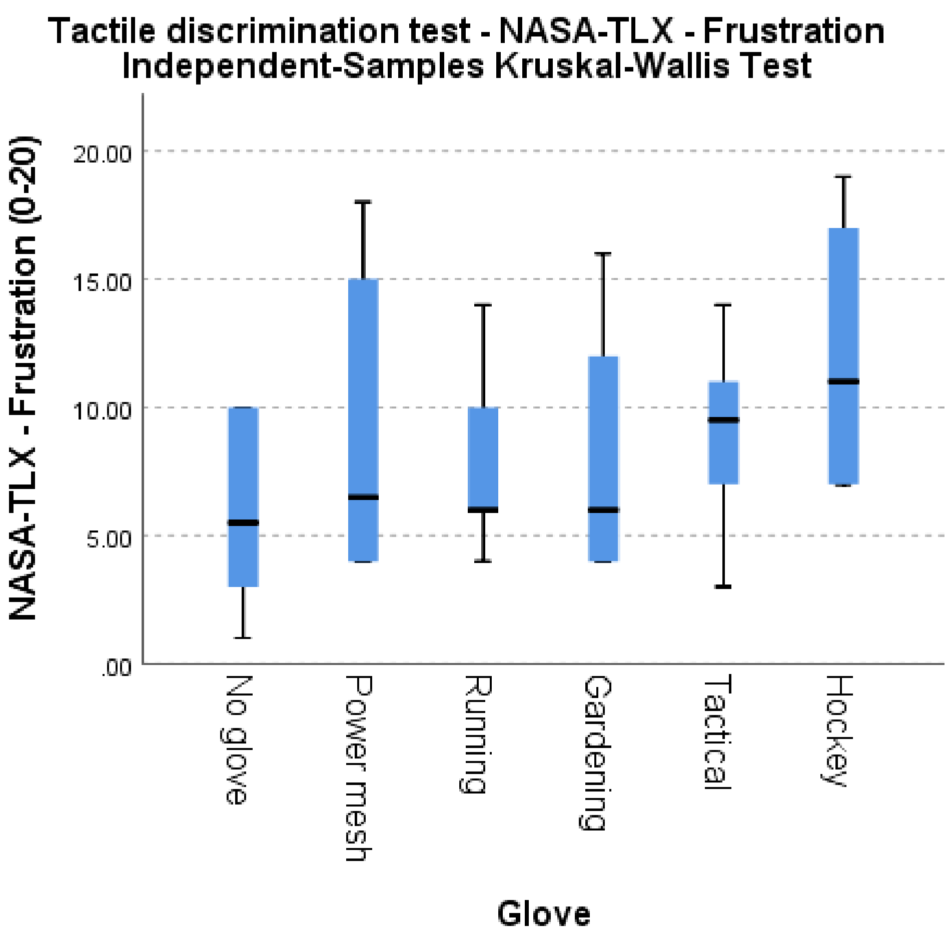
Tactile discrimination test - NASA-TLX - Physical Demand (0-20) field is ordinal but is treated as continuous in the test.

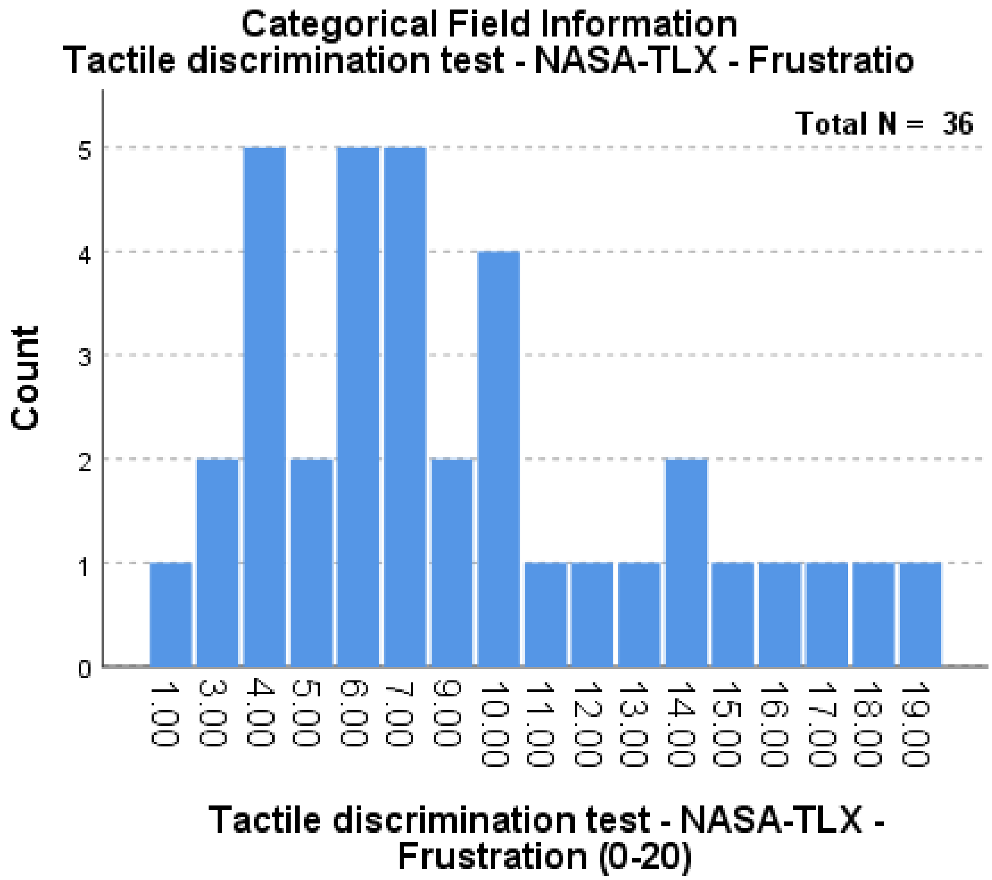
C.3.15 Tactile discrimination test - NASA-TLX - Frustration

Independent-Samples Kruskal-Wallis Test Summary	
Total N	36
Test Statistic	5.711[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.335

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





Tactile discrimination test - NASA-TLX - Frustration (0-20) field is ordinal but is treated as continuous in the test.

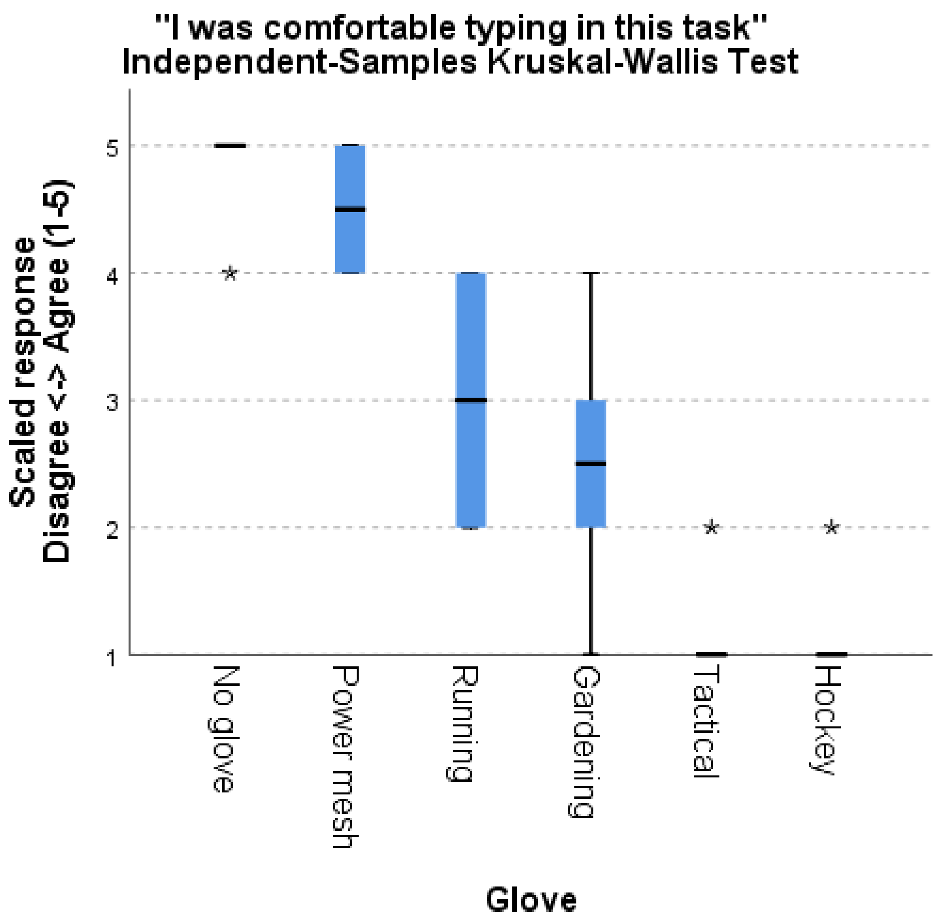
C.4 Scaled response data

C.4.1 "I was comfortable typing in this task"

Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	29.109[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

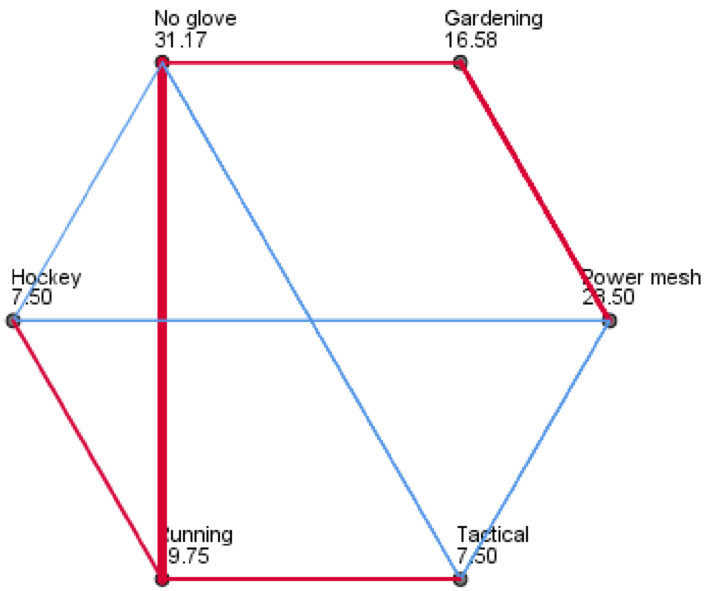
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Tactical-No glove	23.667	5.906	4.007	.000	.001
Hockey-No glove	23.667	5.906	4.007	.000	.001
Tactical-Power mesh	21.000	5.906	3.556	.000	.006
Hockey-Power mesh	21.000	5.906	3.556	.000	.006
Tactical-Running	12.250	5.906	2.074	.038	.571
Hockey-Running	12.250	5.906	2.074	.038	.571
Tactical-Gardening	9.083	5.906	1.538	.124	1.000
Hockey-Gardening	9.083	5.906	1.538	.124	1.000
Tactical-Hockey	.000	5.906	.000	1.000	1.000
Gardening-Running	3.167	5.906	.536	.592	1.000
Gardening-Power mesh	11.917	5.906	2.018	.044	.654
Gardening-No glove	14.583	5.906	2.469	.014	.203
Running-Power mesh	8.750	5.906	1.482	.138	1.000
Running-No glove	11.417	5.906	1.933	.053	.798
Power mesh-No glove	2.667	5.906	.452	.652	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

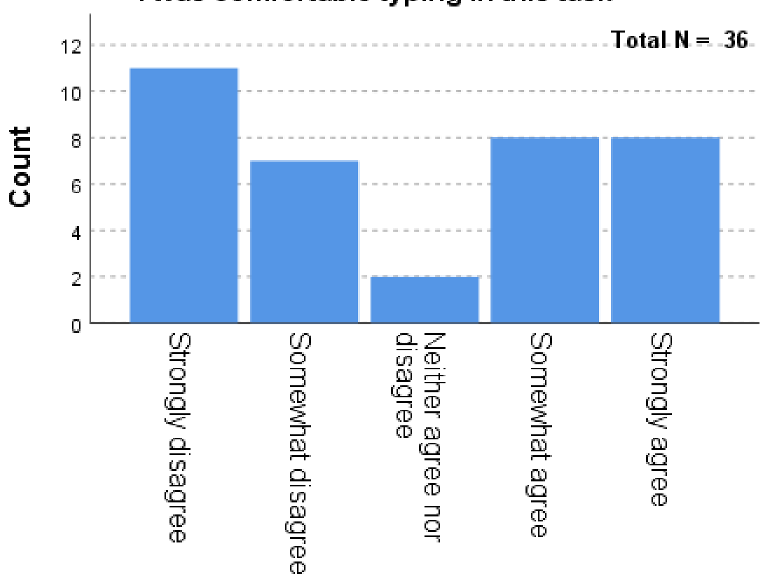
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.

Categorical Field Information
"I was comfortable typing in this task"



"I was comfortable typing in this task"

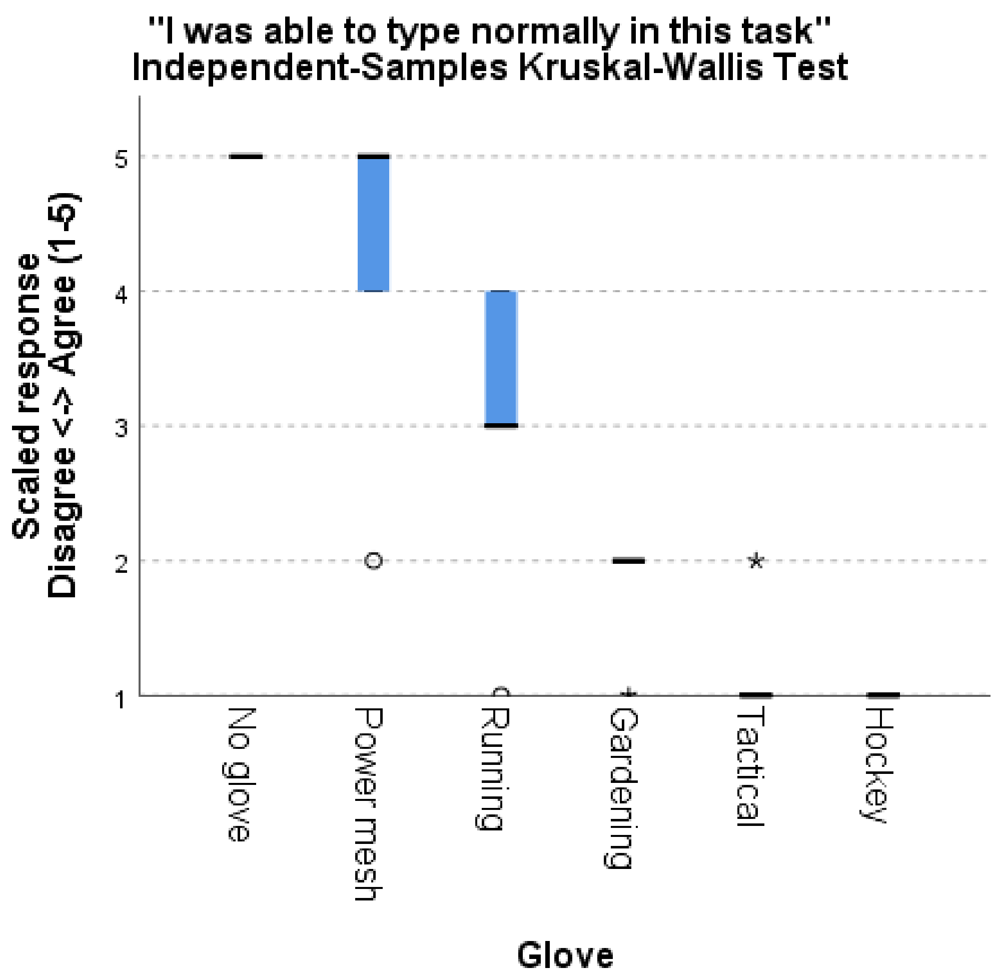
"I was comfortable typing in this task" field is ordinal but is treated as continuous in the test.

C.4.2 "I was able to type normally in this task"

Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	29.390[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



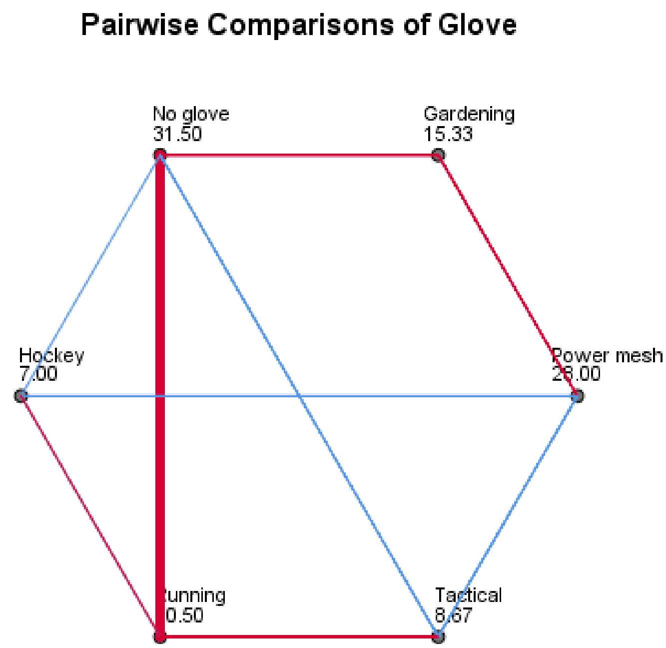
Pairwise Comparisons of Glove

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Hockey-Tactical	1.667	5.846	.285	.776	1.000
Hockey-Gardening	8.333	5.846	1.425	.154	1.000
Hockey-Running	13.500	5.846	2.309	.021	.314
Hockey-Power mesh	21.000	5.846	3.592	.000	.005
Hockey-No glove	24.500	5.846	4.191	.000	.000
Tactical-Gardening	6.667	5.846	1.140	.254	1.000
Tactical-Running	11.833	5.846	2.024	.043	.644
Tactical-Power mesh	19.333	5.846	3.307	.001	.014
Tactical-No glove	22.833	5.846	3.906	.000	.001
Gardening-Running	5.167	5.846	.884	.377	1.000
Gardening-Power mesh	12.667	5.846	2.167	.030	.454
Gardening-No glove	16.167	5.846	2.765	.006	.085
Running-Power mesh	7.500	5.846	1.283	.200	1.000
Running-No glove	11.000	5.846	1.882	.060	.898
Power mesh-No glove	3.500	5.846	.599	.549	1.000

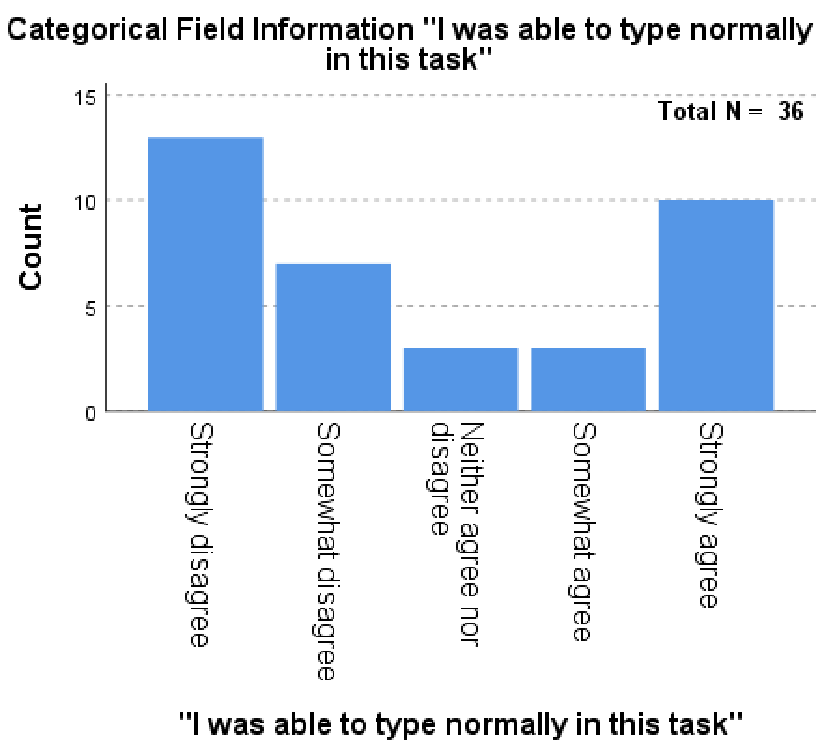
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



Each node shows the sample average rank of Glove.



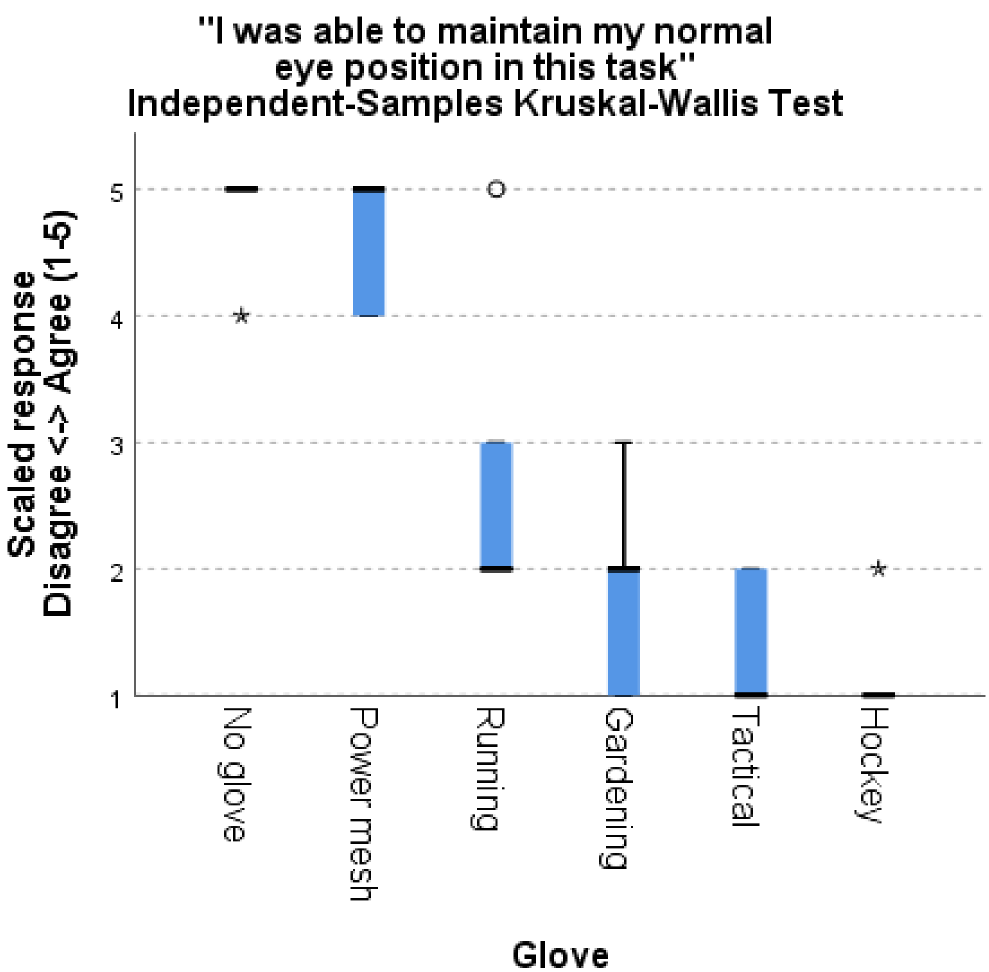
"I was able to type normally in this task" field is ordinal but is treated as continuous in the test.

C.4.3 "I was able to maintain my normal eye position in this task"

Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	27.850[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



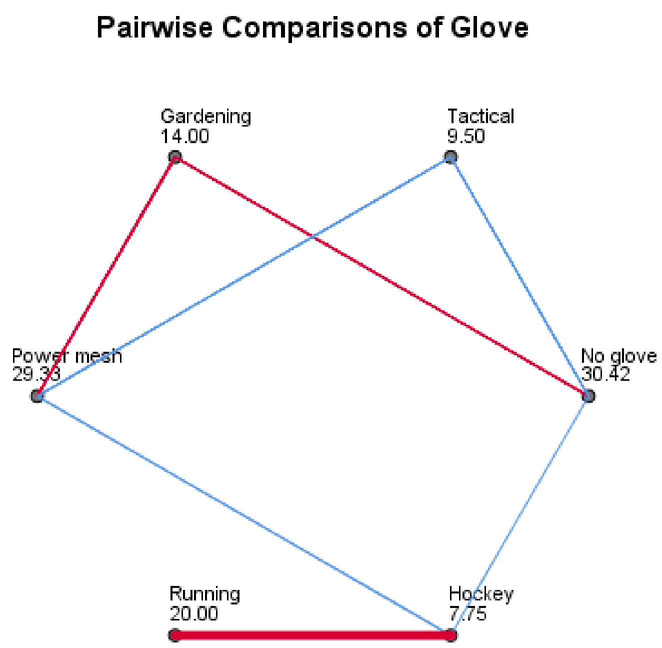
Pairwise Comparisons of Glove

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Hockey-Tactical	1.750	5.861	.299	.765	1.000
Hockey-Gardening	6.250	5.861	1.066	.286	1.000
Hockey-Running	12.250	5.861	2.090	.037	.549
Hockey-Power mesh	21.583	5.861	3.682	.000	.003
Hockey-No glove	22.667	5.861	3.867	.000	.002
Tactical-Gardening	4.500	5.861	.768	.443	1.000
Tactical-Running	10.500	5.861	1.791	.073	1.000
Tactical-Power mesh	19.833	5.861	3.384	.001	.011
Tactical-No glove	20.917	5.861	3.568	.000	.005
Gardening-Running	6.000	5.861	1.024	.306	1.000
Gardening-Power mesh	15.333	5.861	2.616	.009	.133
Gardening-No glove	16.417	5.861	2.801	.005	.076
Running-Power mesh	9.333	5.861	1.592	.111	1.000
Running-No glove	10.417	5.861	1.777	.076	1.000
Power mesh-No glove	1.083	5.861	.185	.853	1.000

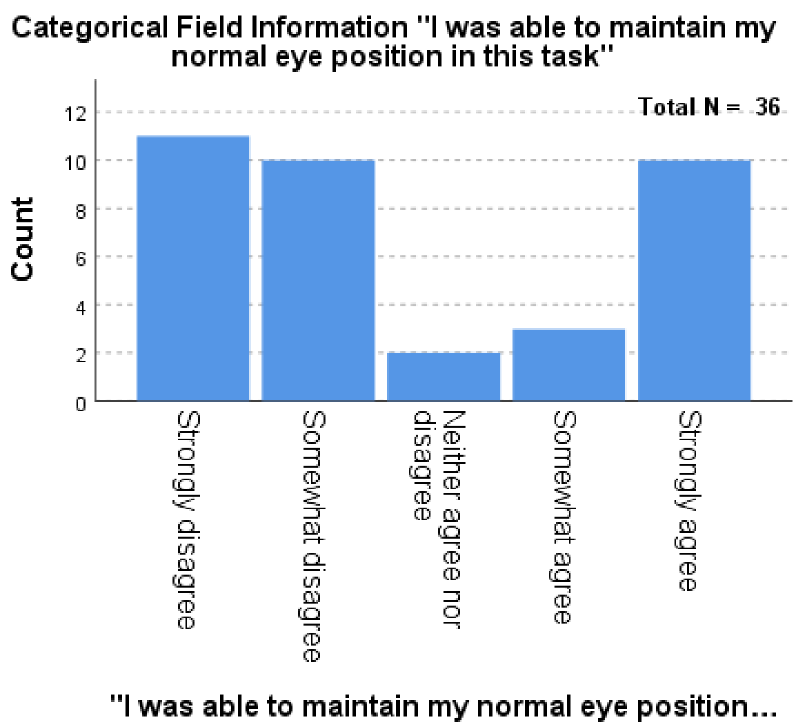
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



Each node shows the sample average rank of Glove.



"I was able to maintain my normal eye position in this task" field is ordinal but is treated as continuous in the test.

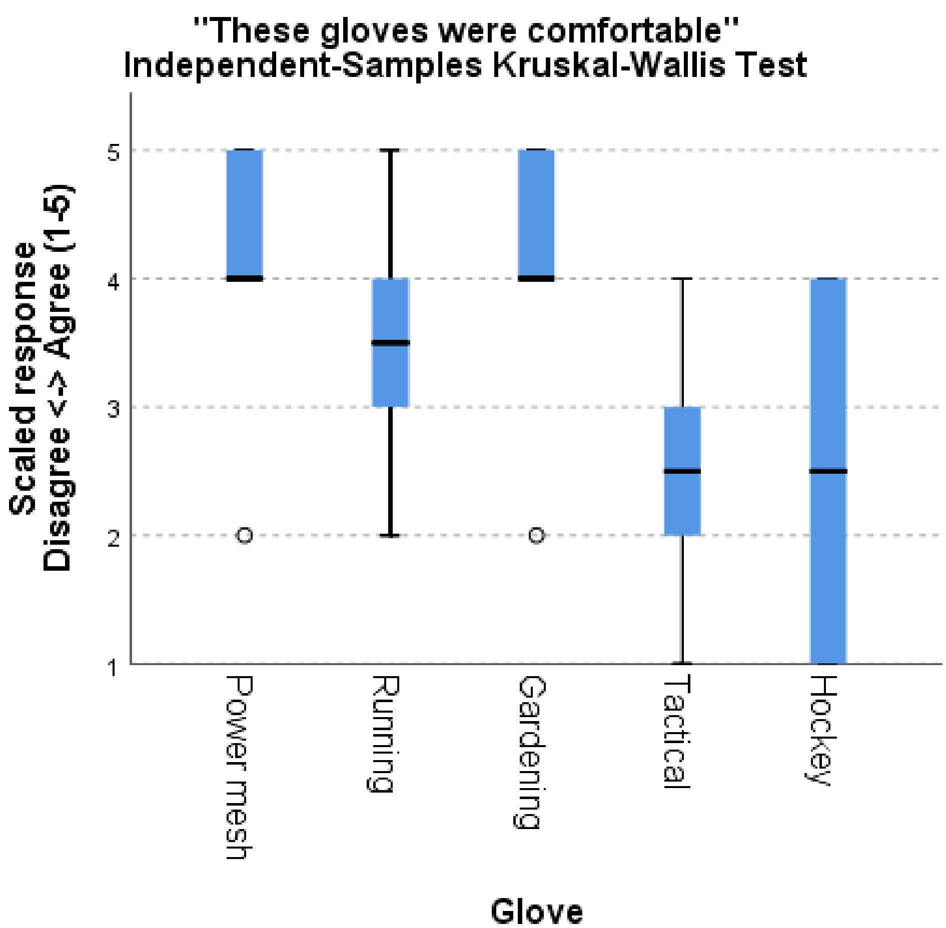
C.4.4 "These gloves were comfortable"

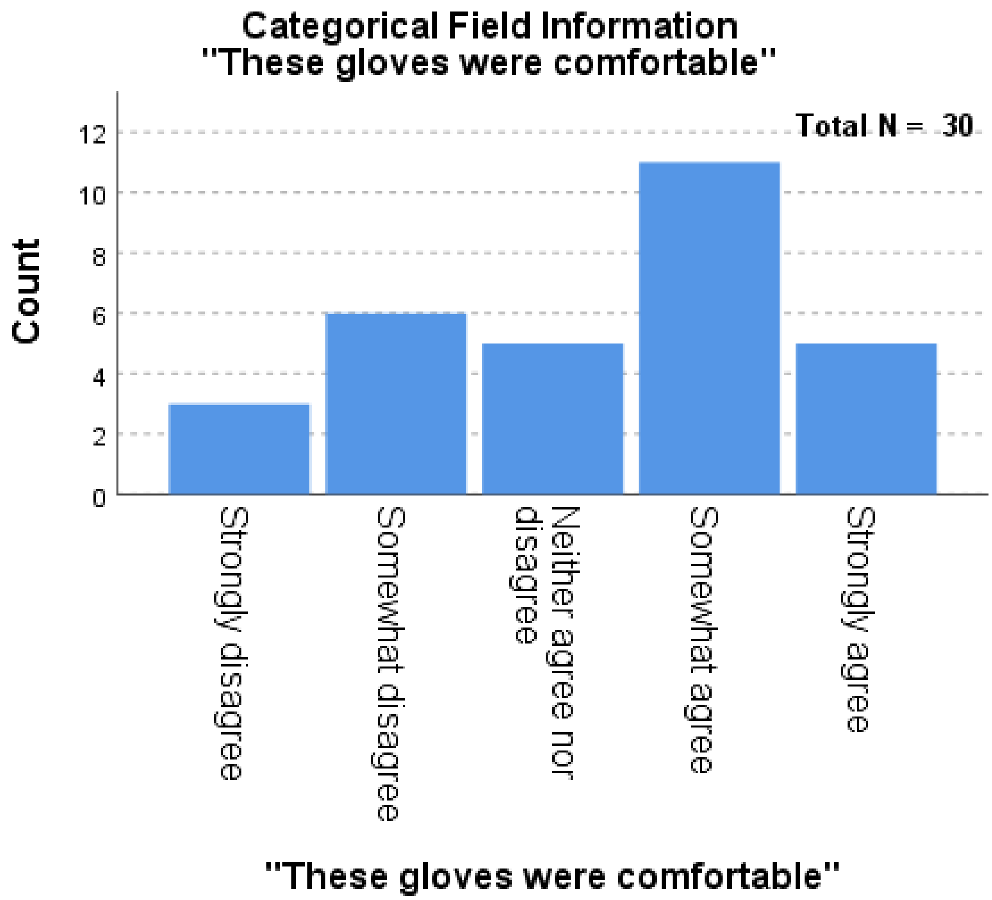
Independent-Samples Kruskal-Wallis Test Summary

Total N	30
Test Statistic	8.886[a,b]
Degree Of Freedom	4
Asymptotic Sig.(2-sided test)	.064

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.



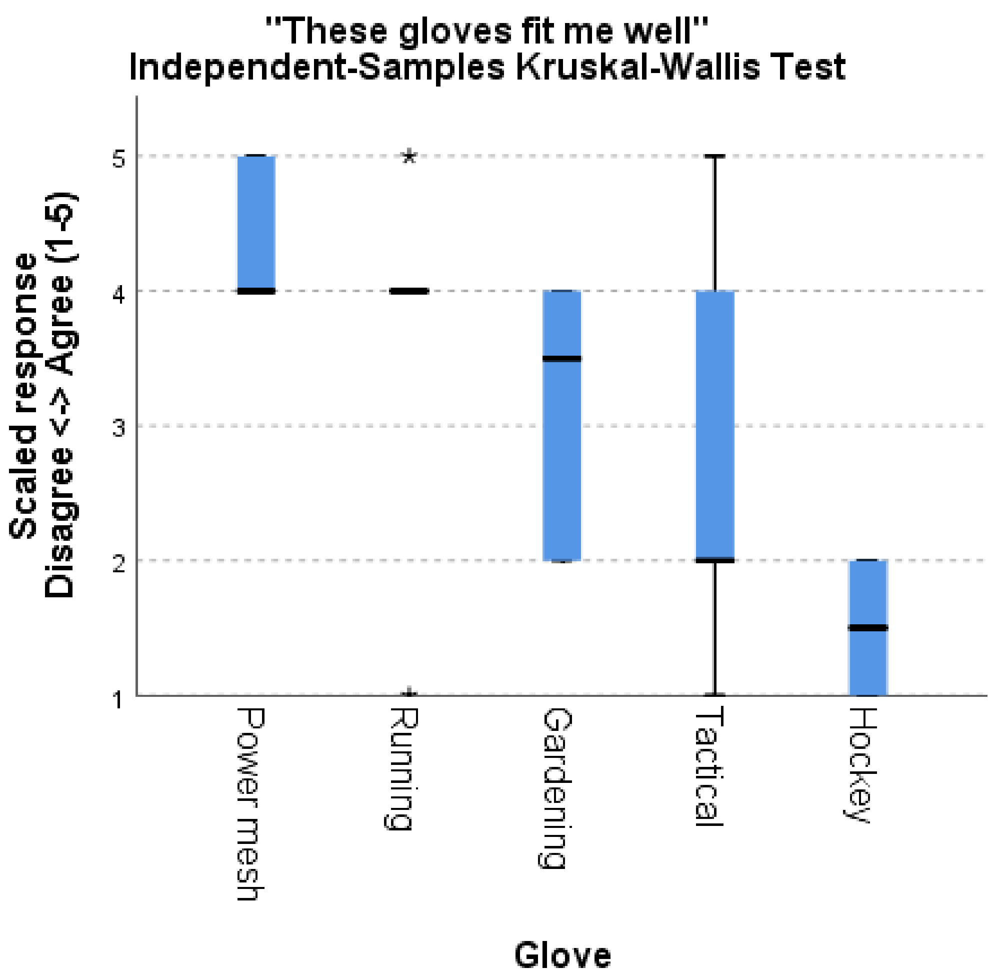


'These gloves were comfortable" field is ordinal but is treated as continuous in the test.

C.4.5 "These gloves fit me well"

Independent-Samples Kruskal-Wallis Test Summary	
Total N	30
Test Statistic	13.632[a]
Degree Of Freedom	4
Asymptotic Sig.(2-sided test)	.009

a. The test statistic is adjusted for ties.



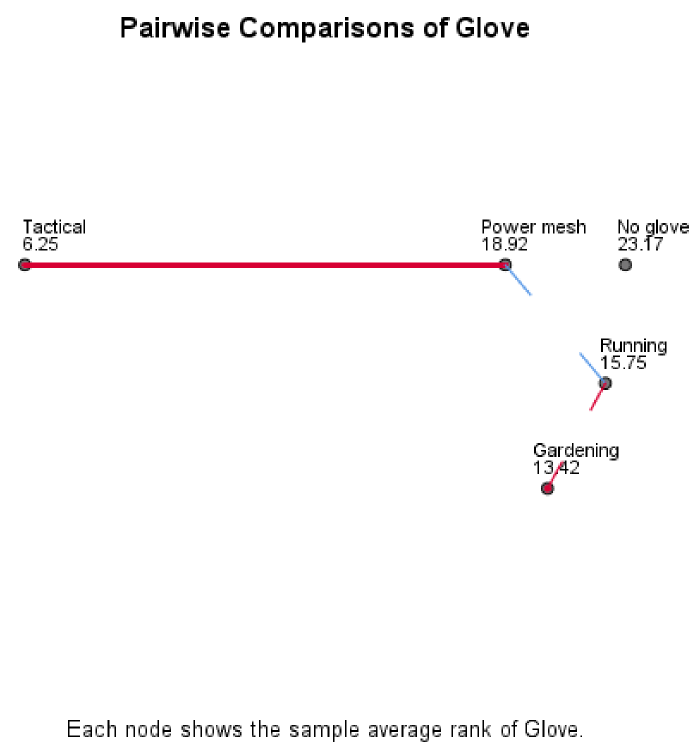
Pairwise Comparisons of Glove

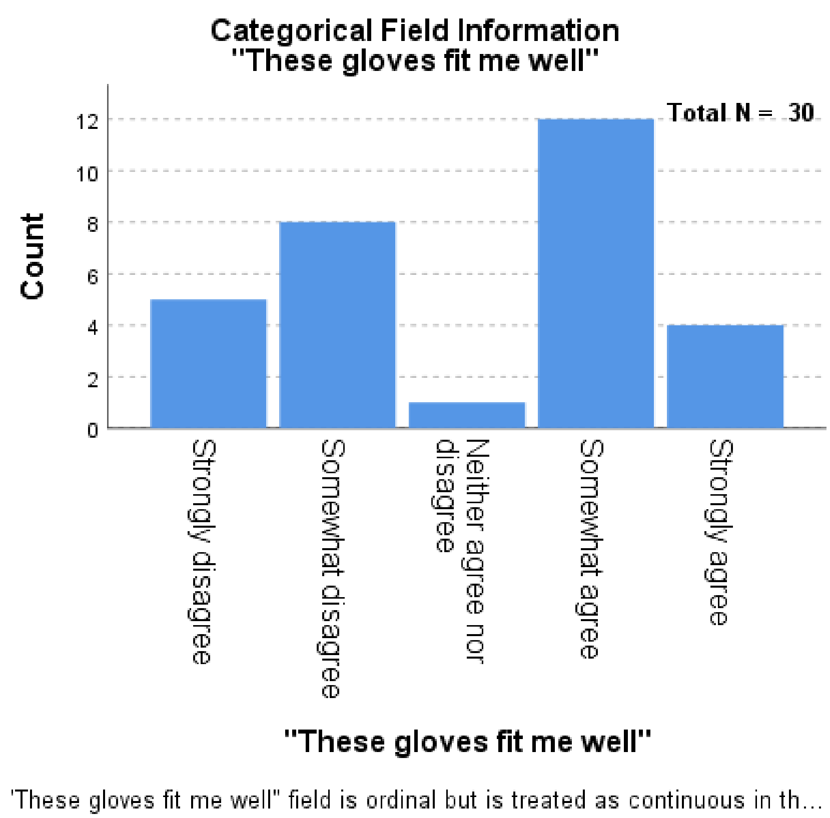
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Hockey-Tactical	7.167	4.851	1.477	.140	1.000
Hockey-Gardening	9.500	4.851	1.958	.050	.753
Hockey-Running	12.667	4.851	2.611	.009	.135
Hockey-Power mesh	16.917	4.851	3.487	.000	.007
Tactical-Gardening	2.333	4.851	.481	.631	1.000
Tactical-Running	5.500	4.851	1.134	.257	1.000
Tactical-Power mesh	9.750	4.851	2.010	.044	.667
Gardening-Running	3.167	4.851	.653	.514	1.000
Gardening-Power mesh	7.417	4.851	1.529	.126	1.000
Running-Power mesh	4.250	4.851	.876	.381	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.


Pairwise Comparisons of Glove
Tactical
6.25
Power mesh
18.92
No glove
23.17
Running
15.75
Gardening
13.42
Each node shows the sample average rank of Glove.

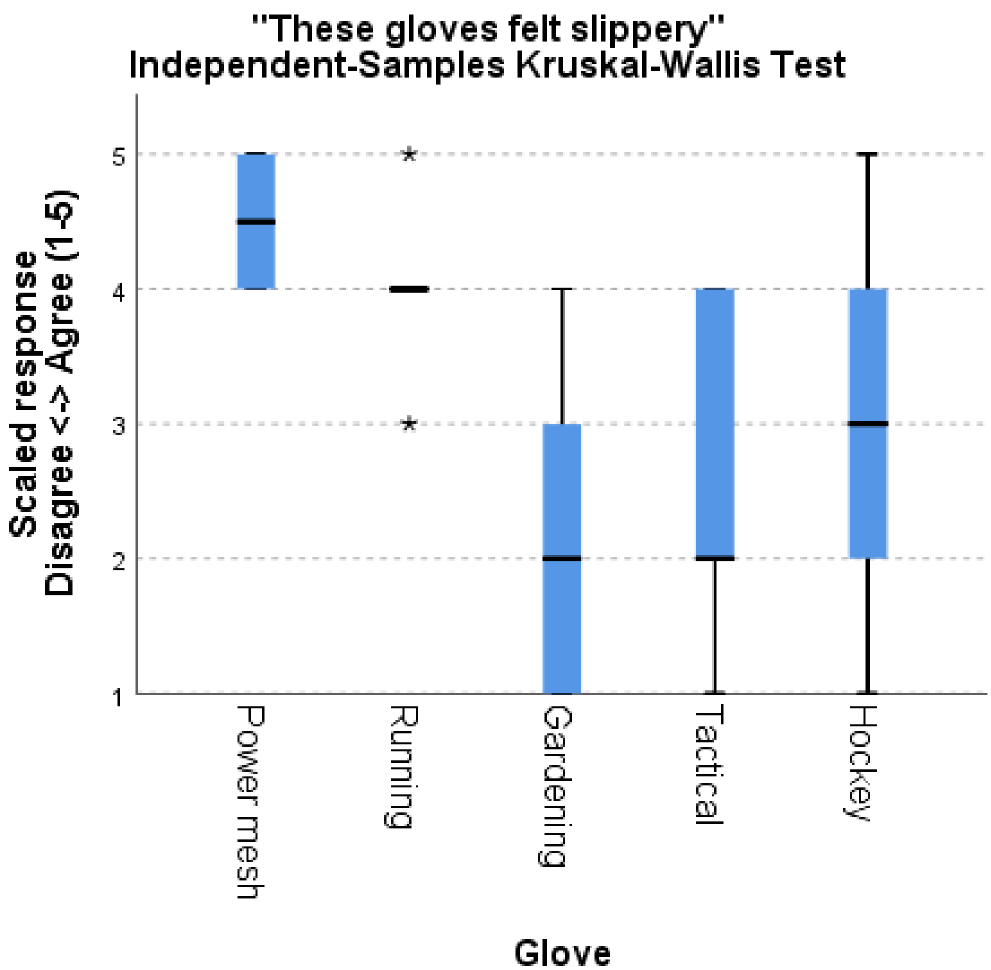

Categorical Field Information
"These gloves fit me well"
Total N = 30
Count
0
2
4
6
8
10
12
Strongly disagree
Somewhat disagree
Neither agree nor disagree
Somewhat agree
Strongly agree
"These gloves fit me well"
'These gloves fit me well" field is ordinal but is treated as continuous in th...

C.4.6 "These gloves felt slippery"

Independent-Samples Kruskal-Wallis Test Summary

Total N	30
Test Statistic	13.537[a]
Degree Of Freedom	4
Asymptotic Sig.(2-sided test)	.009

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

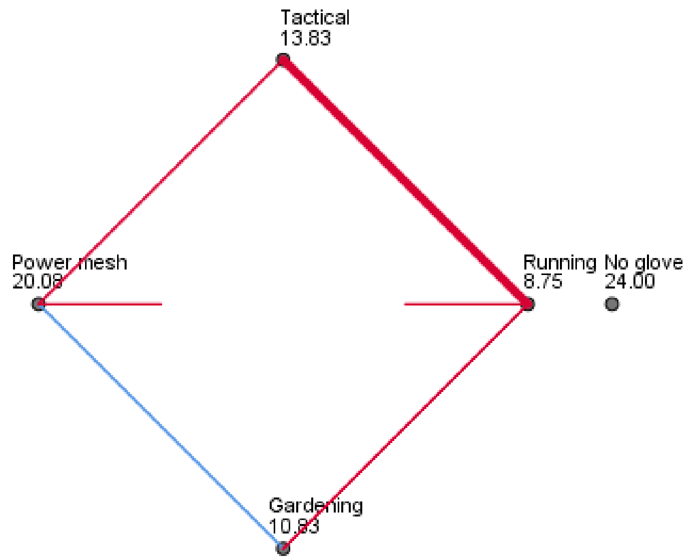
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Gardening-Tactical	-2.083	4.913	-.424	.672	1.000
Gardening-Hockey	-5.083	4.913	-1.035	.301	1.000
Gardening-Running	11.333	4.913	2.307	.021	.316
Gardening-Power mesh	15.250	4.913	3.104	.002	.029
Tactical-Hockey	-3.000	4.913	-.611	.541	1.000
Tactical-Running	9.250	4.913	1.883	.060	.896
Tactical-Power mesh	13.167	4.913	2.680	.007	.110
Hockey-Running	6.250	4.913	1.272	.203	1.000
Hockey-Power mesh	10.167	4.913	2.069	.039	.578
Running-Power mesh	3.917	4.913	.797	.425	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

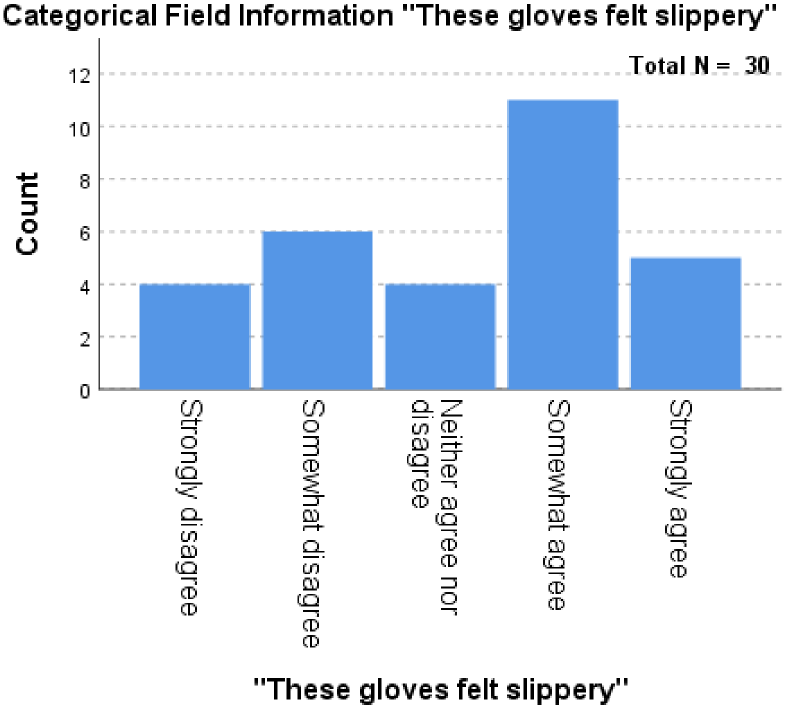
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.

Categorical Field Information "These gloves felt slippery"



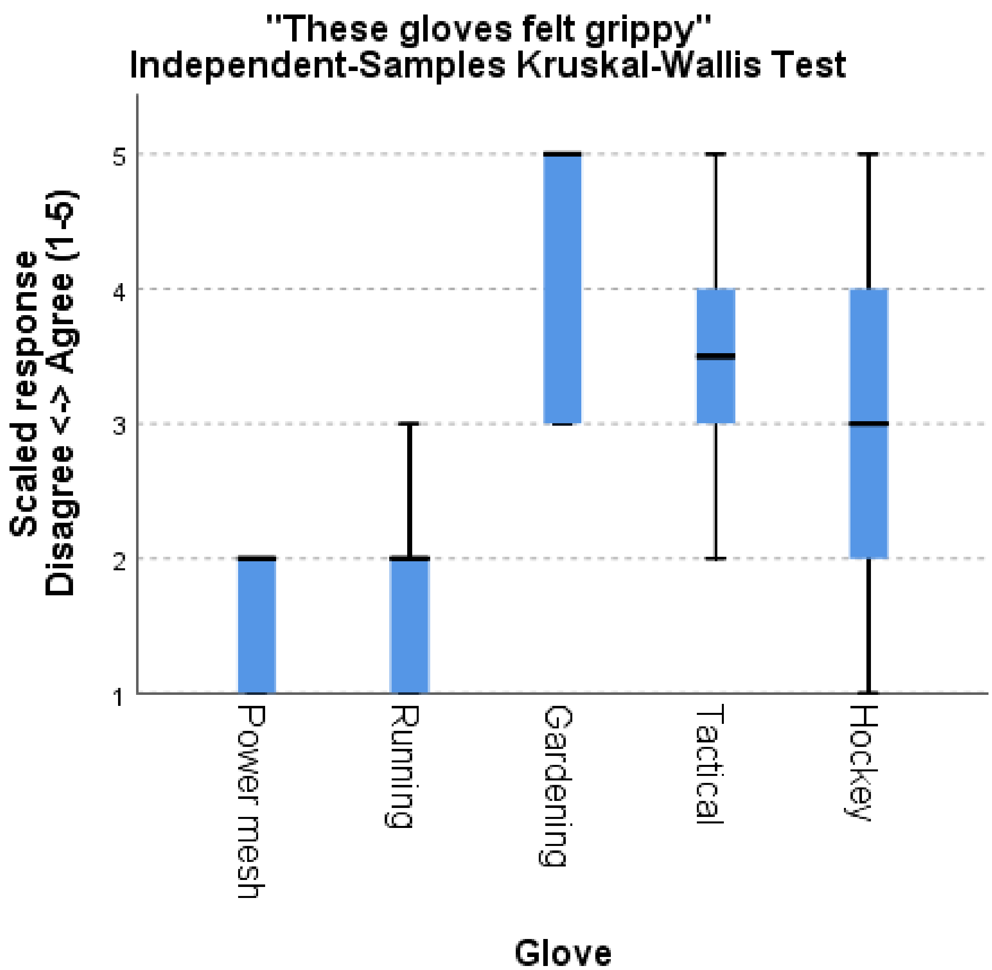
"These gloves felt slippery" field is ordinal but is treated as continuous in ...

C.4.7 "These gloves felt grippy"

Independent-Samples Kruskal-Wallis Test Summary

Total N	30
Test Statistic	15.778[a]
Degree Of Freedom	4
Asymptotic Sig.(2-sided test)	.003

a. The test statistic is adjusted for ties.



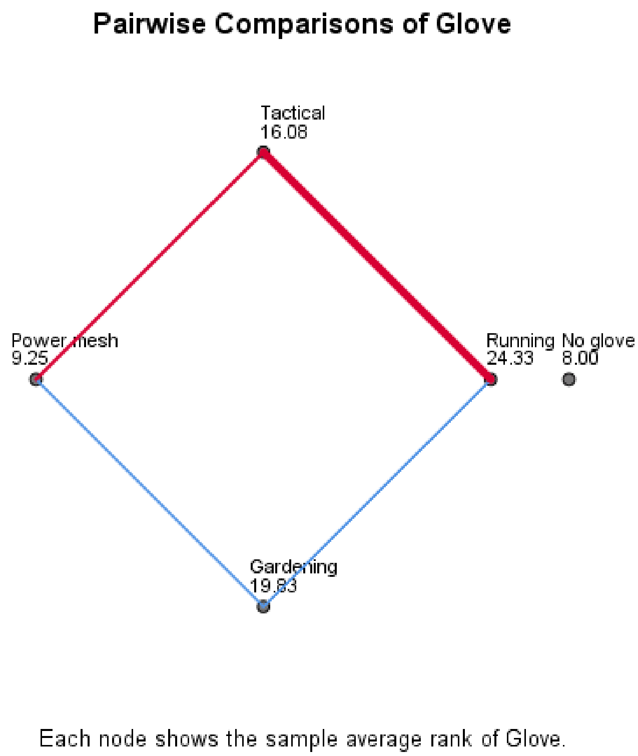
Pairwise Comparisons of Glove

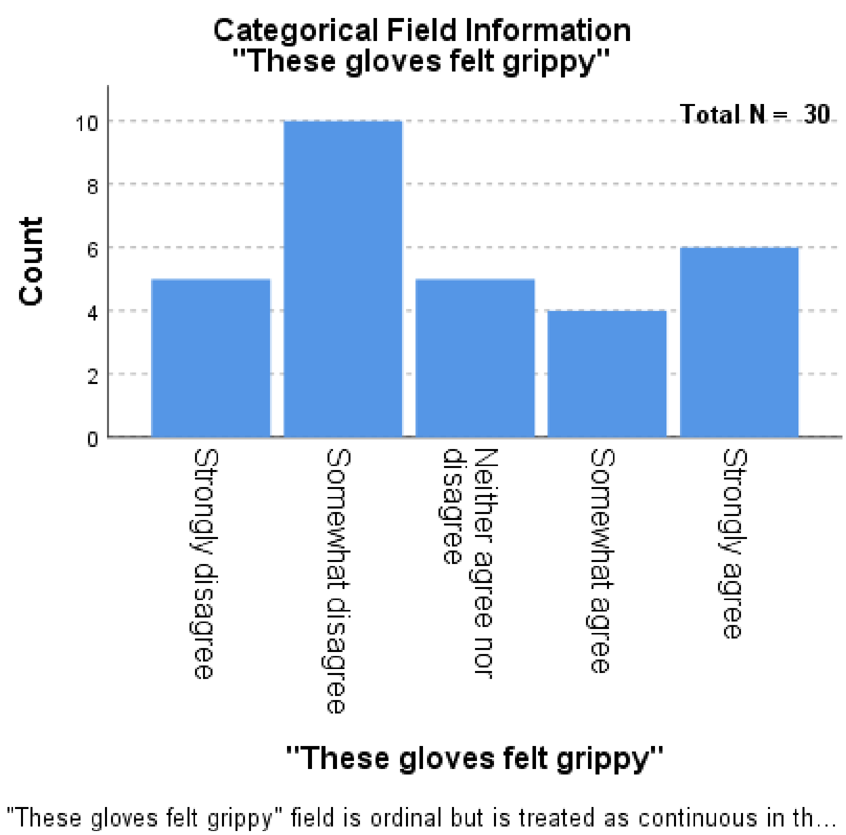
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Power mesh-Running	-1.250	4.939	-.253	.800	1.000
Power mesh-Hockey	-8.083	4.939	-1.637	.102	1.000
Power mesh-Tactical	-11.833	4.939	-2.396	.017	.249
Power mesh-Gardening	-16.333	4.939	-3.307	.001	.014
Running-Hockey	-6.833	4.939	-1.383	.167	1.000
Running-Tactical	-10.583	4.939	-2.143	.032	.482
Running-Gardening	-15.083	4.939	-3.054	.002	.034
Hockey-Tactical	3.750	4.939	.759	.448	1.000
Hockey-Gardening	8.250	4.939	1.670	.095	1.000
Tactical-Gardening	4.500	4.939	.911	.362	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.


Pairwise Comparisons of Glove
Tactical
16.08
Power mesh
9.25
Running
24.33
No glove
8.00
Gardening
19.83
Each node shows the sample average rank of Glove.

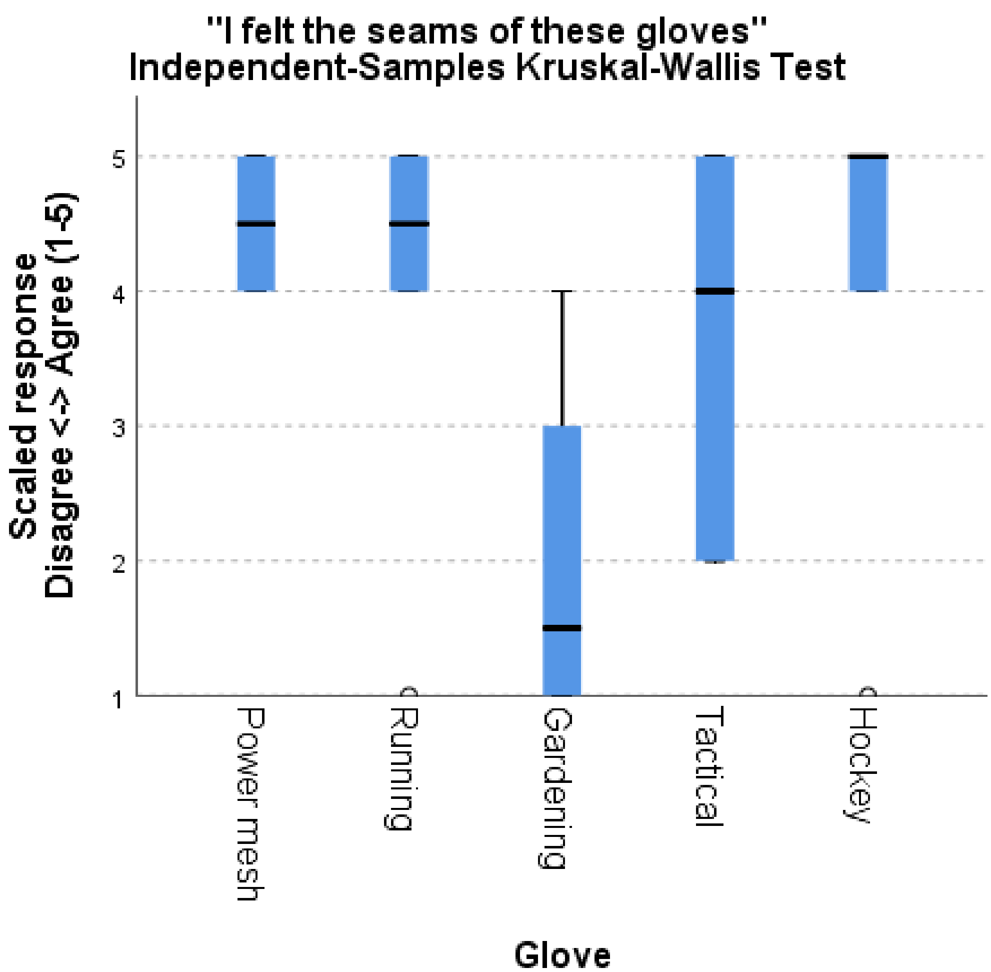

Categorical Field Information
"These gloves felt grippy"
Total N = 30
Count
0
2
4
6
8
10
Strongly disagree
Somewhat disagree
Neither agree nor disagree
Somewhat agree
Strongly agree
"These gloves felt grippy"
"These gloves felt grippy" field is ordinal but is treated as continuous in th...

C.4.8 "I felt the seams of these gloves"

Independent-Samples Kruskal-Wallis Test Summary

Total N	30
Test Statistic	9.613[a]
Degree Of Freedom	4
Asymptotic Sig.(2-sided test)	.047

a. The test statistic is adjusted for ties.



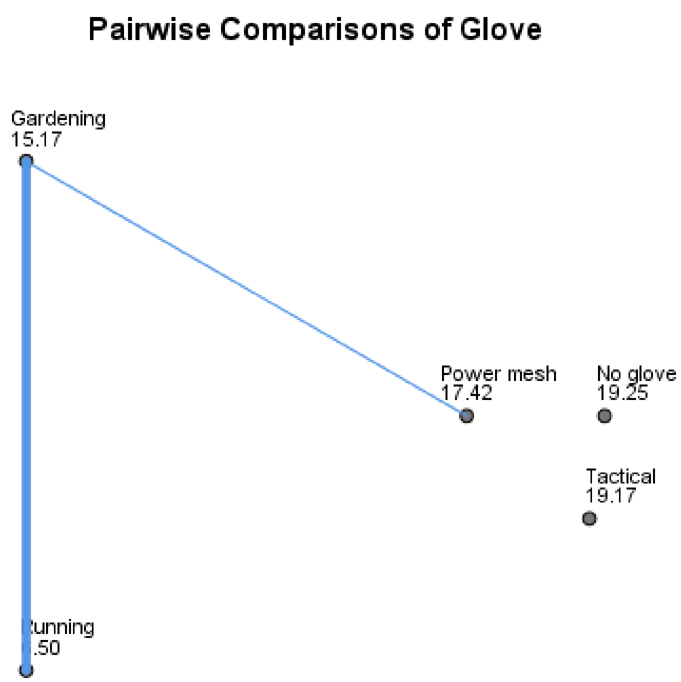
Pairwise Comparisons of Glove

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Gardening-Tactical	-8.667	4.833	-1.793	.073	1.000
Gardening-Running	10.917	4.833	2.259	.024	.359
Gardening-Hockey	-12.667	4.833	-2.621	.009	.132
Gardening-Power mesh	12.750	4.833	2.638	.008	.125
Tactical-Running	2.250	4.833	.466	.642	1.000
Tactical-Hockey	-4.000	4.833	-.828	.408	1.000
Tactical-Power mesh	4.083	4.833	.845	.398	1.000
Running-Hockey	-1.750	4.833	-.362	.717	1.000
Running-Power mesh	1.833	4.833	.379	.704	1.000
Hockey-Power mesh	.083	4.833	.017	.986	1.000

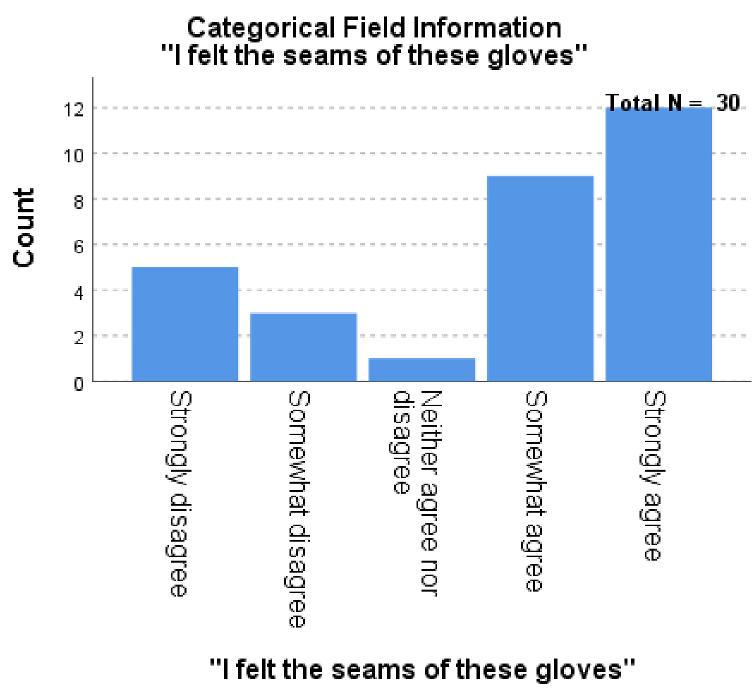
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



Each node shows the sample average rank of Glove.



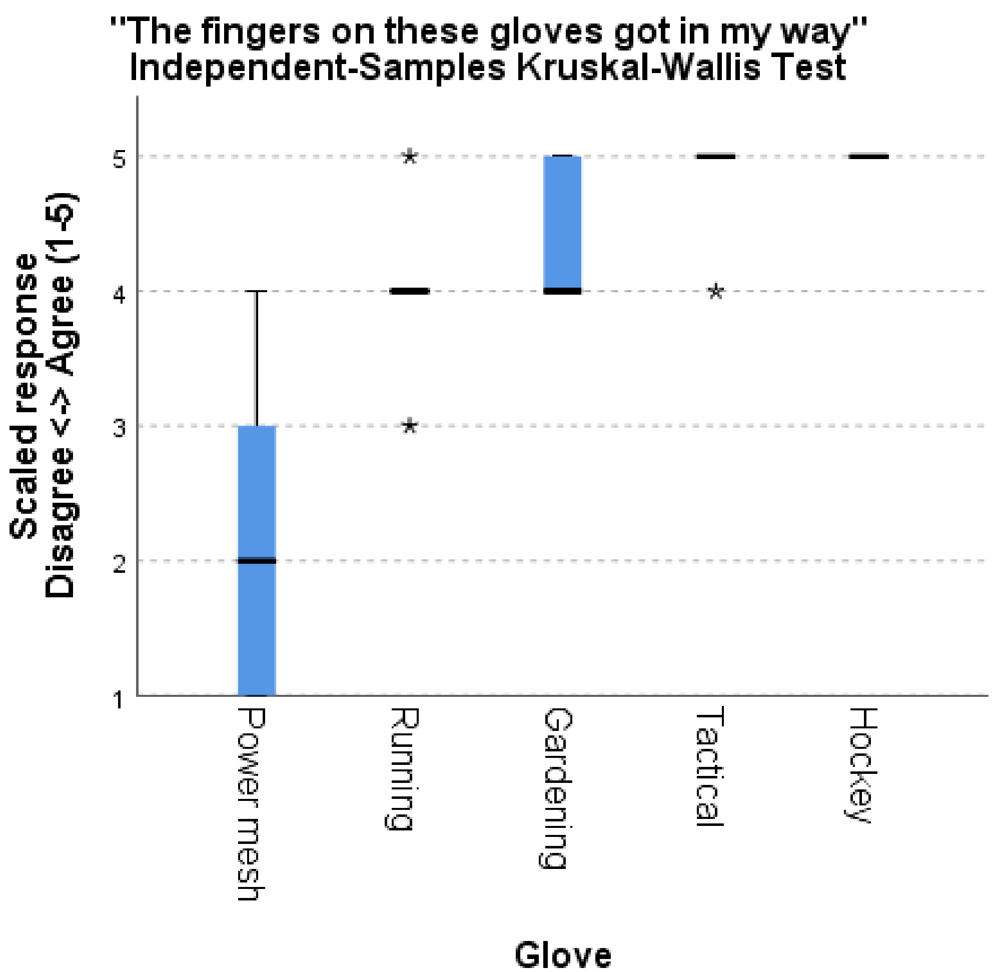
"I felt the seams of these gloves" field is ordinal but is treated as continuous in the test.

C.4.9 "The fingers on these gloves got in my way" across Glove

Independent-Samples Kruskal-Wallis Test Summary

Total N	30
Test Statistic	20.671[a]
Degree Of Freedom	4
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

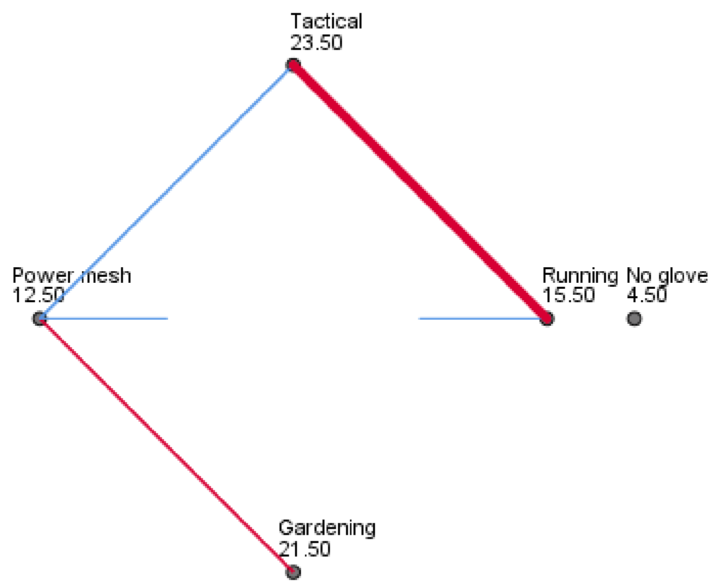
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
Power mesh-Running	-8.000	4.717	-1.696	.090	1.000
Power mesh-Gardening	-11.000	4.717	-2.332	.020	.296
Power mesh-Tactical	-17.000	4.717	-3.604	.000	.005
Power mesh-Hockey	-19.000	4.717	-4.028	.000	.001
Running-Gardening	-3.000	4.717	-.636	.525	1.000
Running-Tactical	-9.000	4.717	-1.908	.056	.846
Running-Hockey	-11.000	4.717	-2.332	.020	.296
Gardening-Tactical	-6.000	4.717	-1.272	.203	1.000
Gardening-Hockey	-8.000	4.717	-1.696	.090	1.000
Tactical-Hockey	-2.000	4.717	-.424	.672	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

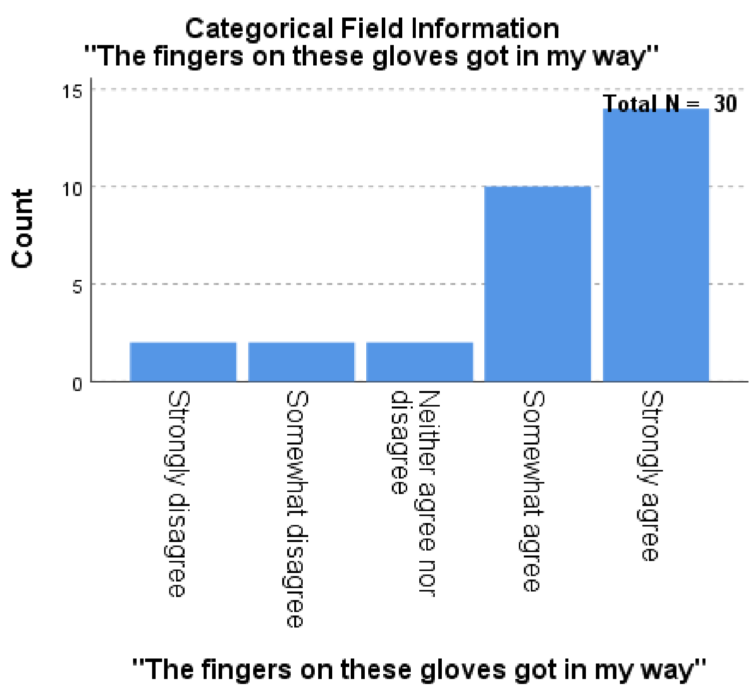
Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.



"The fingers on these gloves got in my way" field is ordinal but is treated as continuous in the test.

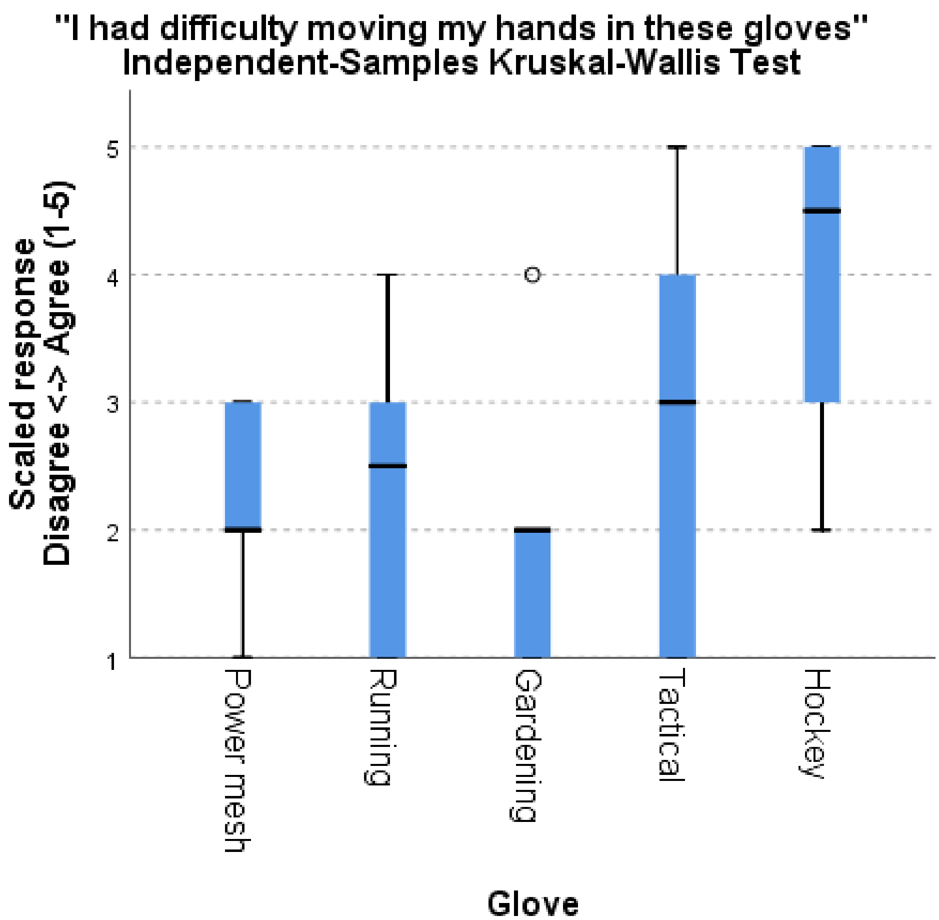
C.4.10 "I had difficulty moving my hands in these gloves"

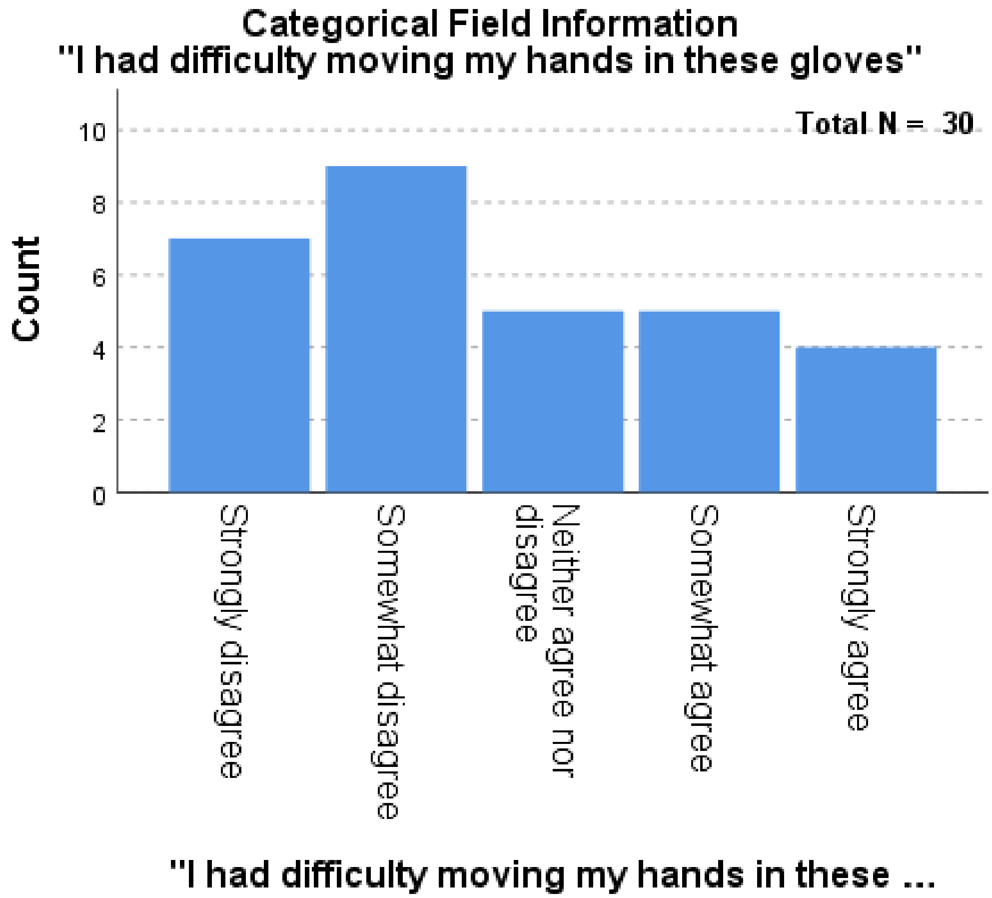
Independent-Samples Kruskal-Wallis Test Summary

Total N	30
Test Statistic	7.335[a,b]
Degree Of Freedom	4
Asymptotic Sig.(2-sided test)	.119

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





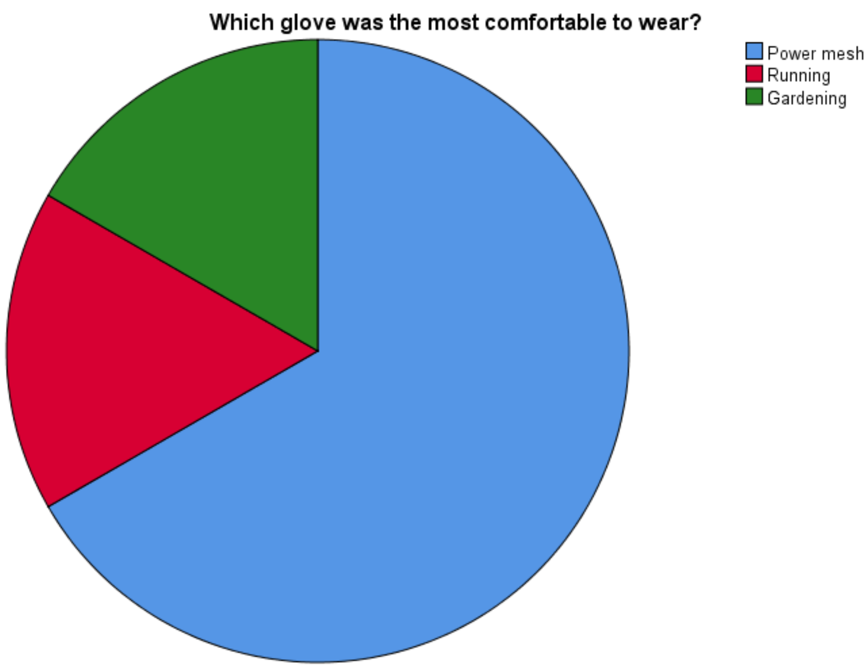
"I had difficulty moving my hands in these gloves" field is ordinal but is treated as continuous in the test.

C.5 Glove preference data

C.5.1 Which glove was the most comfortable to wear?

Which glove was the most comfortable to wear?

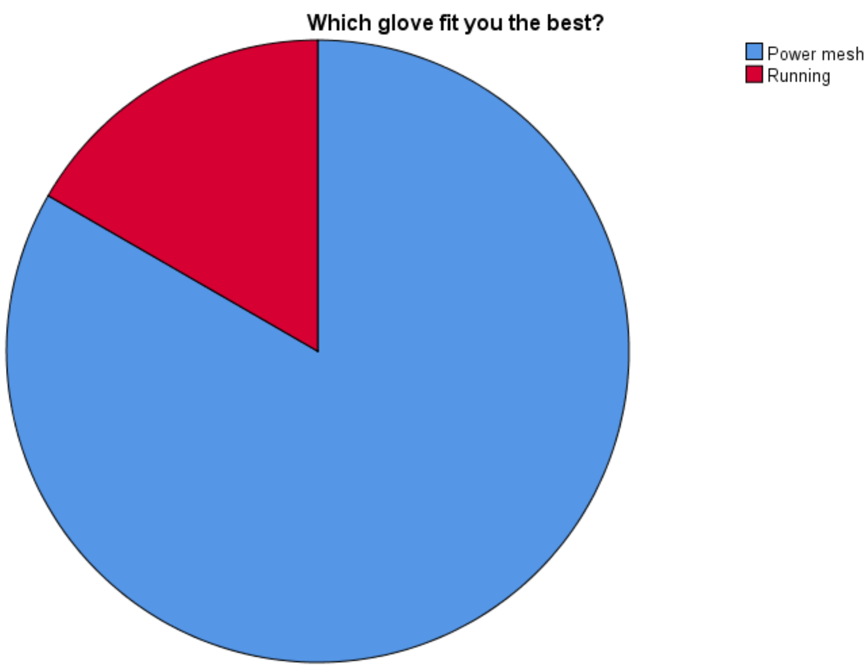
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Power mesh	4	66.7	66.7	66.7
	Running	1	16.7	16.7	83.3
	Gardening	1	16.7	16.7	100.0
	Total	6	100.0	100.0	



C.5.2 Which glove fit you the best?

Which glove fit you the best?

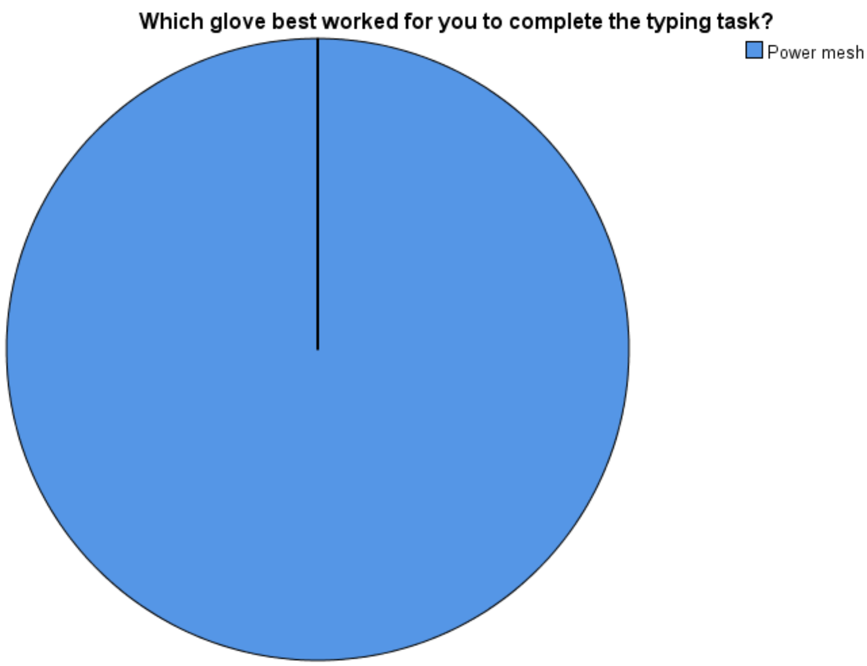
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Power mesh	5	83.3	83.3	83.3
	Running	1	16.7	16.7	100.0
	Total	6	100.0	100.0	



C.5.3 Which glove best worked for you to complete the typing task?

Which glove best worked for you to complete the typing task?

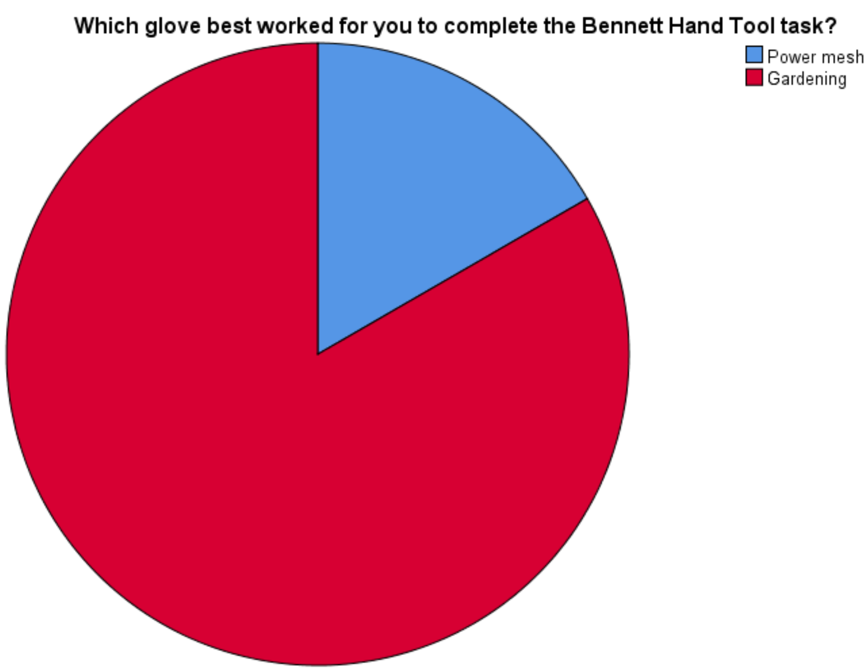
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Power mesh	6	100.0	100.0	100.0



C.5.4 Which glove best worked for you to complete the Bennett Hand Tool task?

Which glove best worked for you to complete the Bennett Hand Tool task?

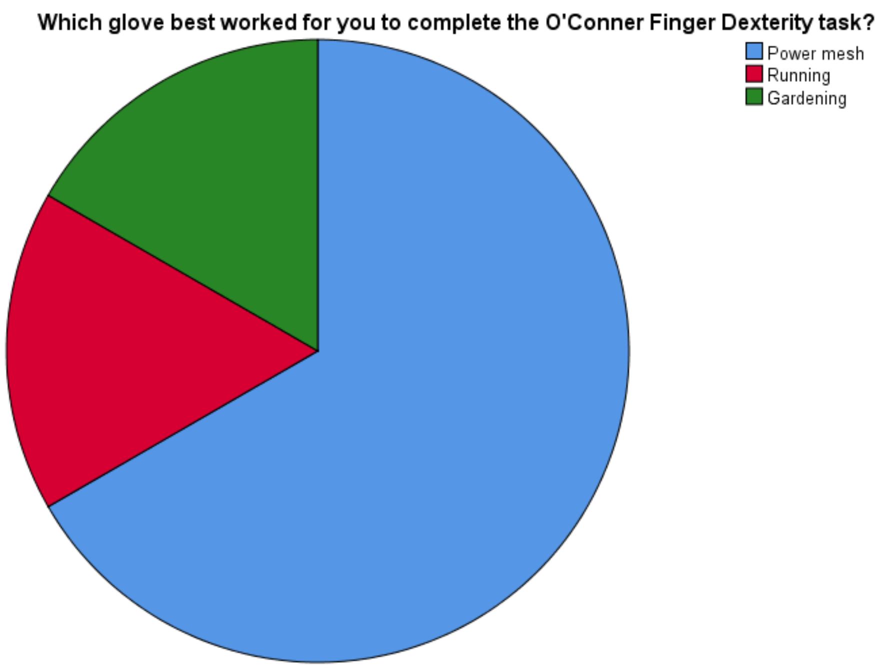
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Power mesh	1	16.7	16.7	16.7
	Gardening	5	83.3	83.3	100.0
	Total	6	100.0	100.0	



C.5.5 Which glove best worked for you to complete the O'Conner Finger Dexterity task?

Which glove best worked for you to complete the O'Conner Finger Dexterity task?

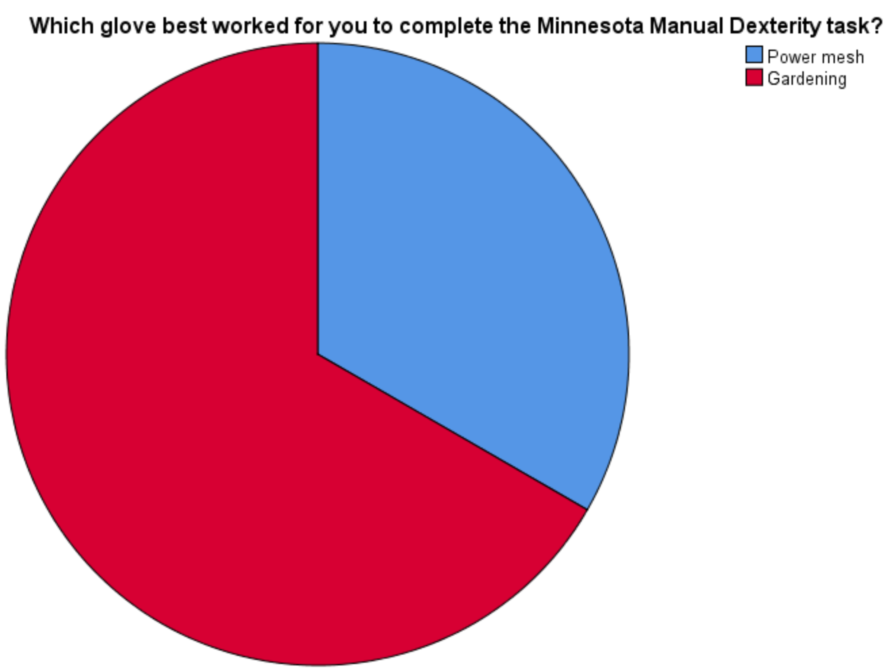
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Power mesh	4	66.7	66.7	66.7
	Running	1	16.7	16.7	83.3
	Gardening	1	16.7	16.7	100.0
	Total	6	100.0	100.0	



C.5.6 Which glove best worked for you to complete the Minnesota Manual Dexterity task?

Which glove best worked for you to complete the Minnesota Manual Dexterity task?

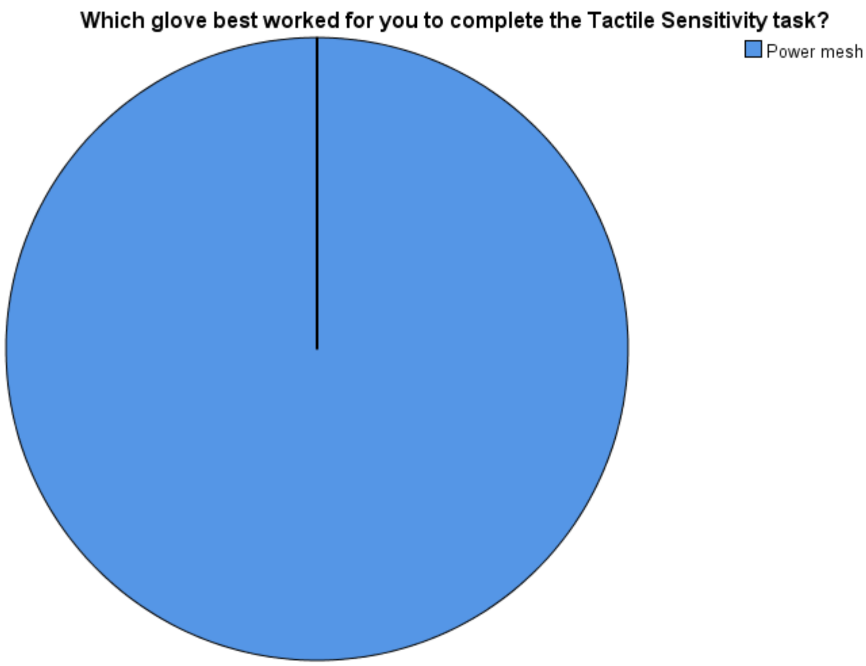
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Power mesh	2	33.3	33.3	33.3
	Gardening	4	66.7	66.7	100.0
	Total	6	100.0	100.0	



C.5.7 Which glove best worked for you to complete the Tactile Sensitivity task?

Which glove best worked for you to complete the Tactile Sensitivity task?

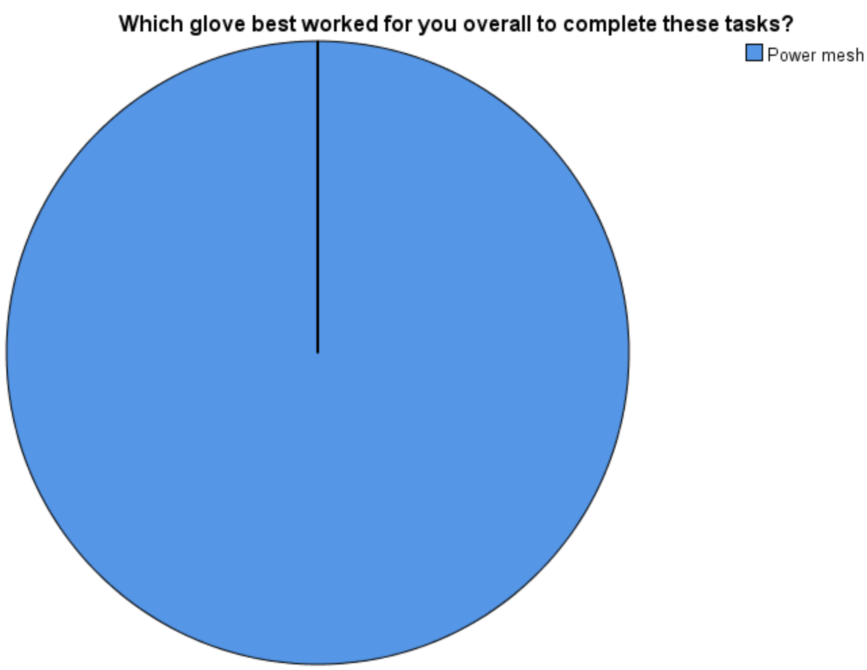
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Power mesh	6	100.0	100.0	100.0



C.5.8 Which glove best worked for you overall to complete these tasks?

Which glove best worked for you overall to complete these tasks?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Power mesh	6	100.0	100.0	100.0



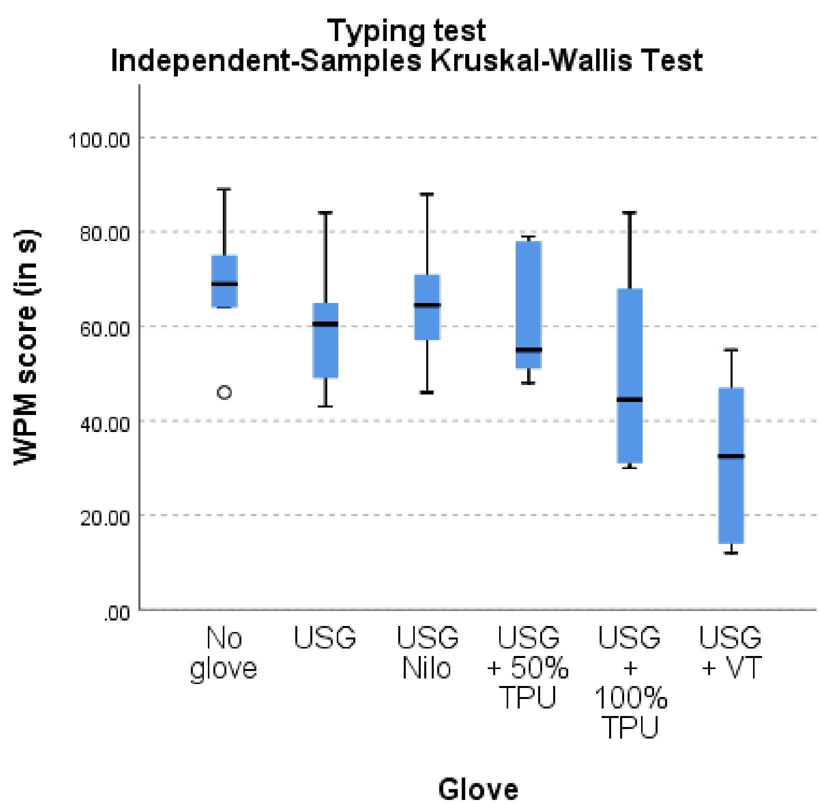
APPENDIX D. STUDY 3: CONSTRUCTED GLOVE DATA

D.1 Task performance data

D.1.1 Typing test

Independent-Samples Kruskal-Wallis Test Summary	
Total N	36
Test Statistic	12.304[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.031

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

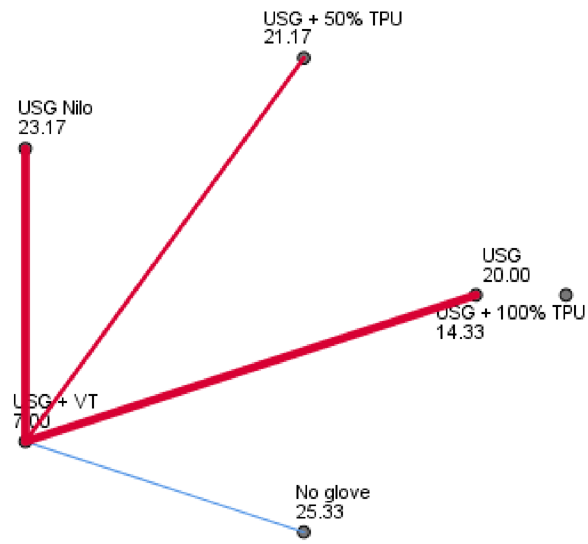
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
USG + VT-USG + 100% TPU	7.333	6.080	1.206	.228	1.000
USG + VT-USG	13.000	6.080	2.138	.033	.488
USG + VT-USG + 50% TPU	14.167	6.080	2.330	.020	.297
USG + VT-USG Nilo	16.167	6.080	2.659	.008	.118
USG + VT-No glove	18.333	6.080	3.015	.003	.039
USG + 100% TPU-USG	5.667	6.080	.932	.351	1.000
USG + 100% TPU-USG + 50% TPU	6.833	6.080	1.124	.261	1.000
USG + 100% TPU-USG Nilo	8.833	6.080	1.453	.146	1.000
USG + 100% TPU-No glove	11.000	6.080	1.809	.070	1.000
USG-USG + 50% TPU	-1.167	6.080	-.192	.848	1.000
USG-USG Nilo	-3.167	6.080	-.521	.603	1.000
USG-No glove	5.333	6.080	.877	.380	1.000
USG + 50% TPU-USG Nilo	2.000	6.080	.329	.742	1.000
USG + 50% TPU-No glove	4.167	6.080	.685	.493	1.000
USG Nilo-No glove	2.167	6.080	.356	.722	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

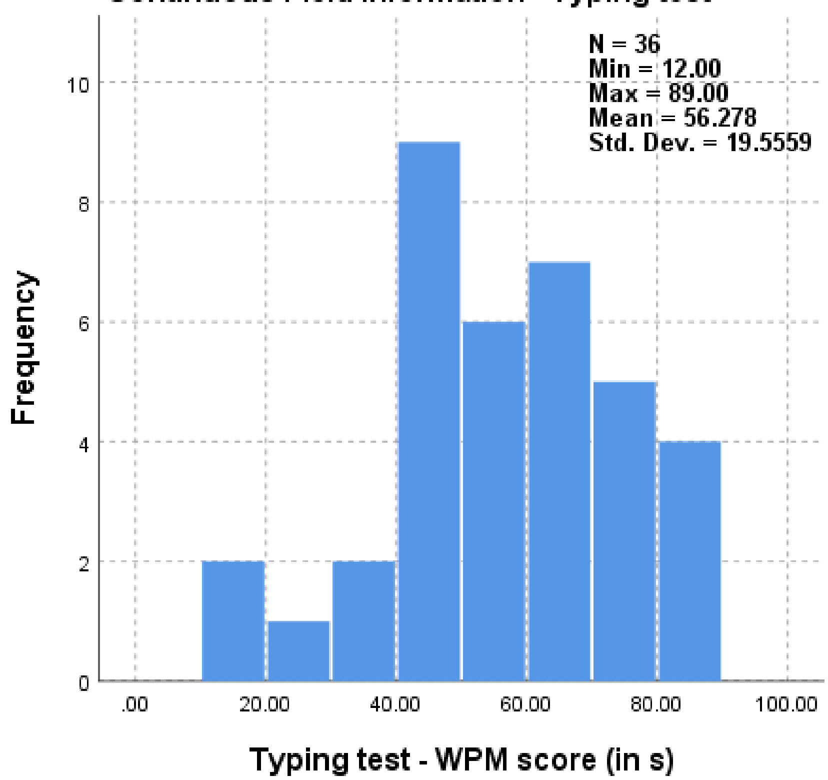
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.

Continuous Field Information - Typing test

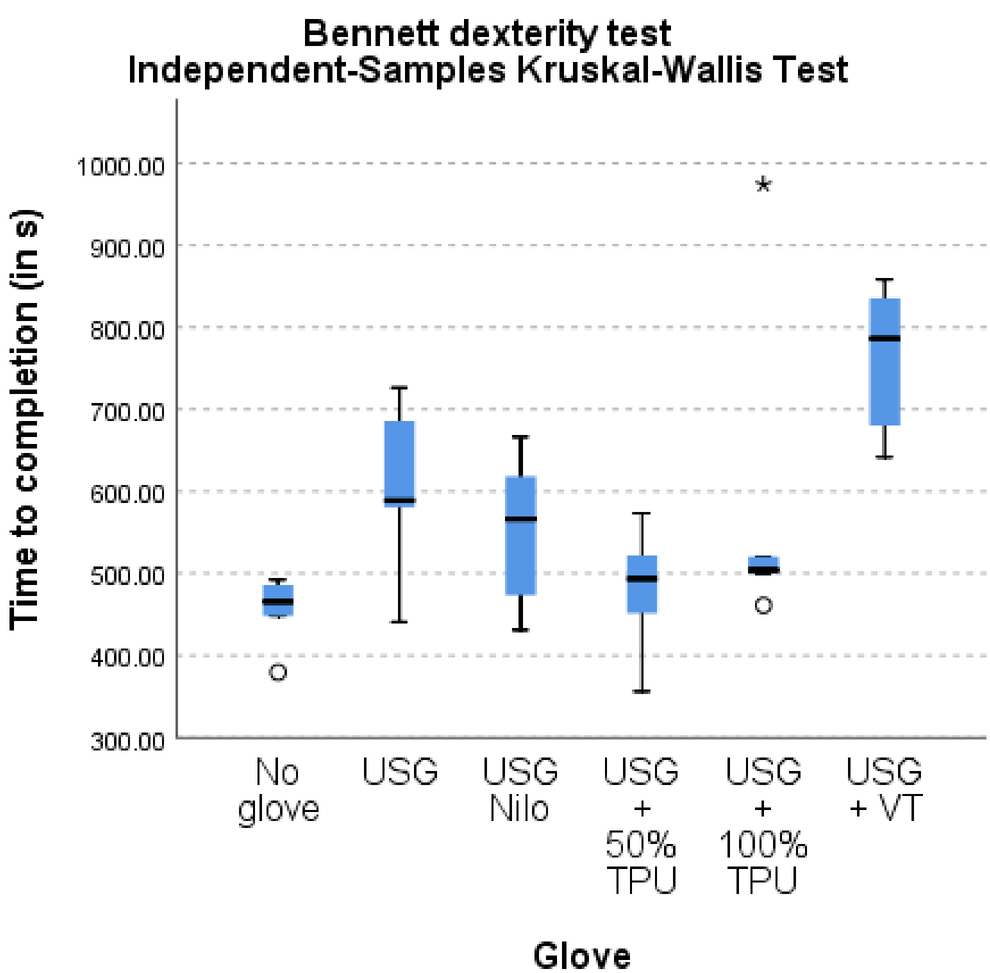


D.1.2 Bennett dexterity test

Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	18.372[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.003

a. The test statistic is adjusted for ties.



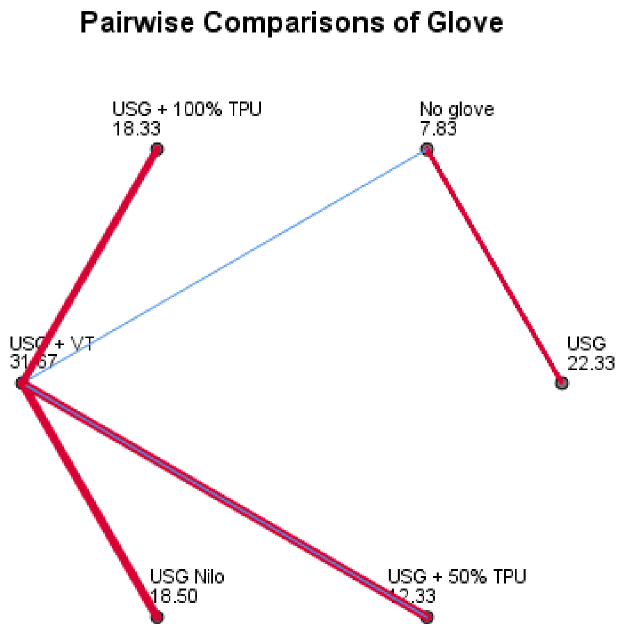
Pairwise Comparisons of Glove

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
No glove-USG + 50% TPU	-4.500	6.083	-.740	.459	1.000
No glove-USG + 100% TPU	-10.500	6.083	-1.726	.084	1.000
No glove-USG Nilo	-10.667	6.083	-1.754	.080	1.000
No glove-USG	-14.500	6.083	-2.384	.017	.257
No glove-USG + VT	-23.833	6.083	-3.918	.000	.001
USG + 50% TPU-USG + 100% TPU	-6.000	6.083	-.986	.324	1.000
USG + 50% TPU-USG Nilo	6.167	6.083	1.014	.311	1.000
USG + 50% TPU-USG	10.000	6.083	1.644	.100	1.000
USG + 50% TPU-USG + VT	-19.333	6.083	-3.178	.001	.022
USG + 100% TPU-USG Nilo	.167	6.083	.027	.978	1.000
USG + 100% TPU-USG	4.000	6.083	.658	.511	1.000
USG + 100% TPU-USG + VT	-13.333	6.083	-2.192	.028	.426
USG Nilo-USG	3.833	6.083	.630	.529	1.000
USG Nilo-USG + VT	-13.167	6.083	-2.165	.030	.456
USG-USG + VT	-9.333	6.083	-1.534	.125	1.000

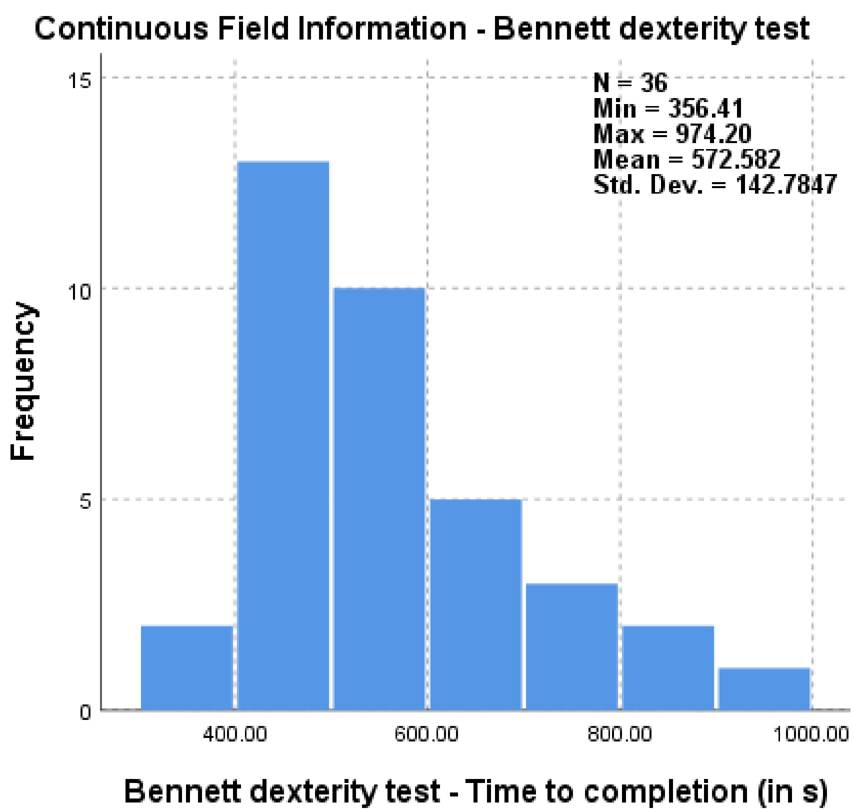
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



Each node shows the sample average rank of Glove.

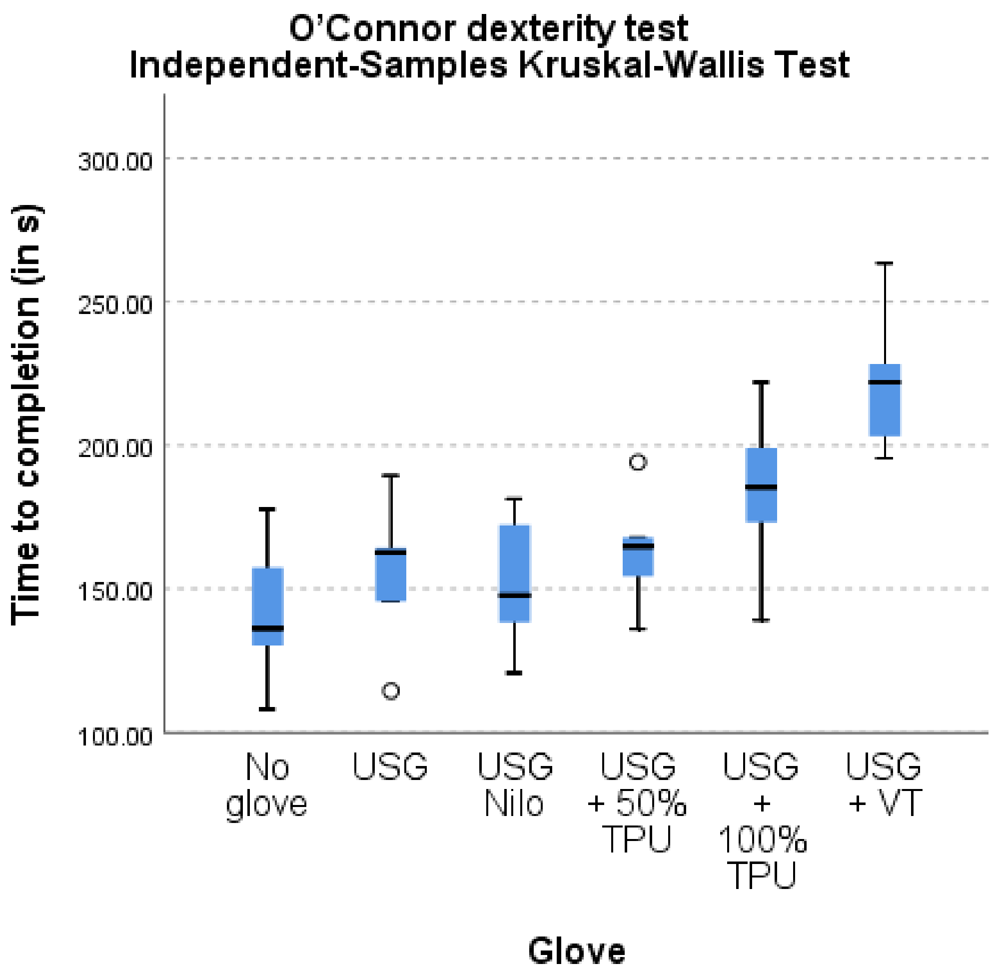


D.1.3 O'Connor dexterity test

Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	19.871[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.001

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

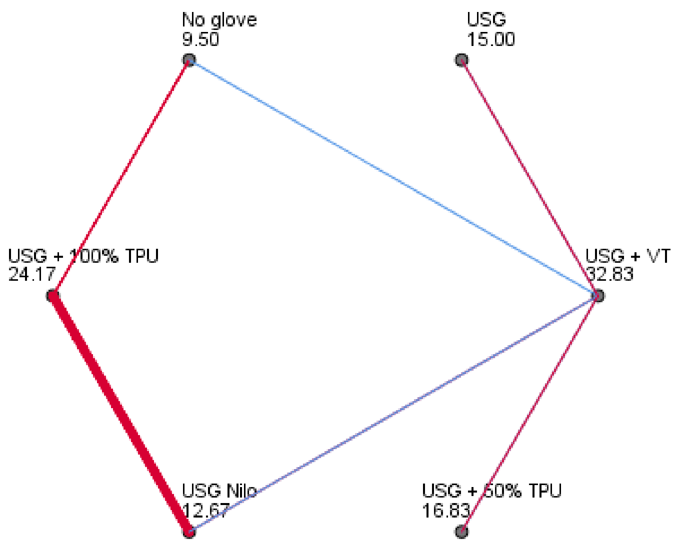
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
No glove-USG Nilo	-3.167	6.083	-.521	.603	1.000
No glove-USG	-5.500	6.083	-.904	.366	1.000
No glove-USG + 50% TPU	-7.333	6.083	-1.206	.228	1.000
No glove-USG + 100% TPU	-14.667	6.083	-2.411	.016	.239
No glove-USG + VT	-23.333	6.083	-3.836	.000	.002
USG Nilo-USG	2.333	6.083	.384	.701	1.000
USG Nilo-USG + 50% TPU	-4.167	6.083	-.685	.493	1.000
USG Nilo-USG + 100% TPU	-11.500	6.083	-1.891	.059	.880
USG Nilo-USG + VT	-20.167	6.083	-3.315	.001	.014
USG-USG + 50% TPU	-1.833	6.083	-.301	.763	1.000
USG-USG + 100% TPU	-9.167	6.083	-1.507	.132	1.000
USG-USG + VT	-17.833	6.083	-2.932	.003	.051
USG + 50% TPU-USG + 100% TPU	-7.333	6.083	-1.206	.228	1.000
USG + 50% TPU-USG + VT	-16.000	6.083	-2.630	.009	.128
USG + 100% TPU-USG + VT	-8.667	6.083	-1.425	.154	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

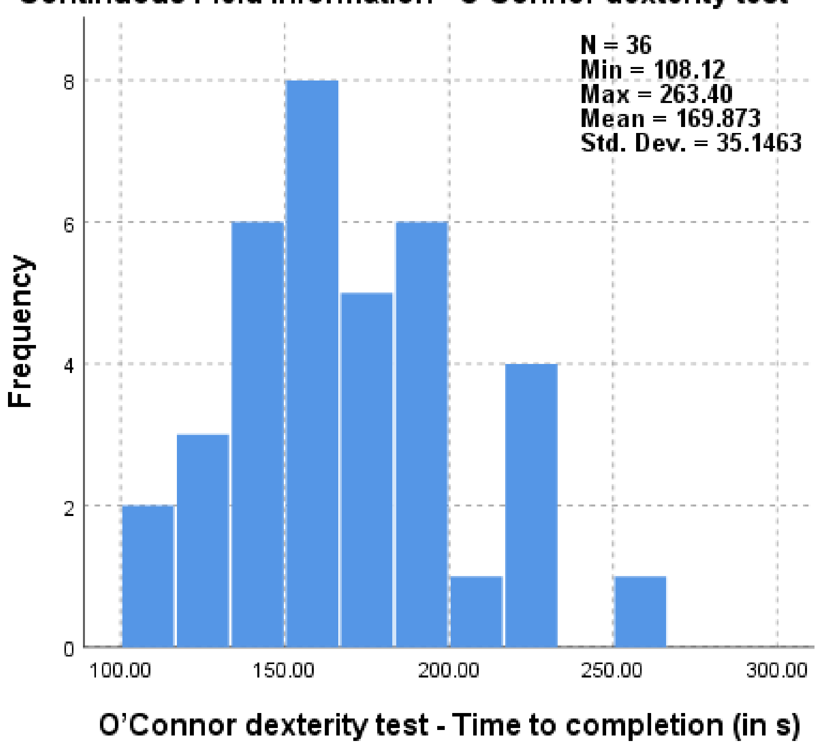
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.

Continuous Field Information - O'Connor dexterity test

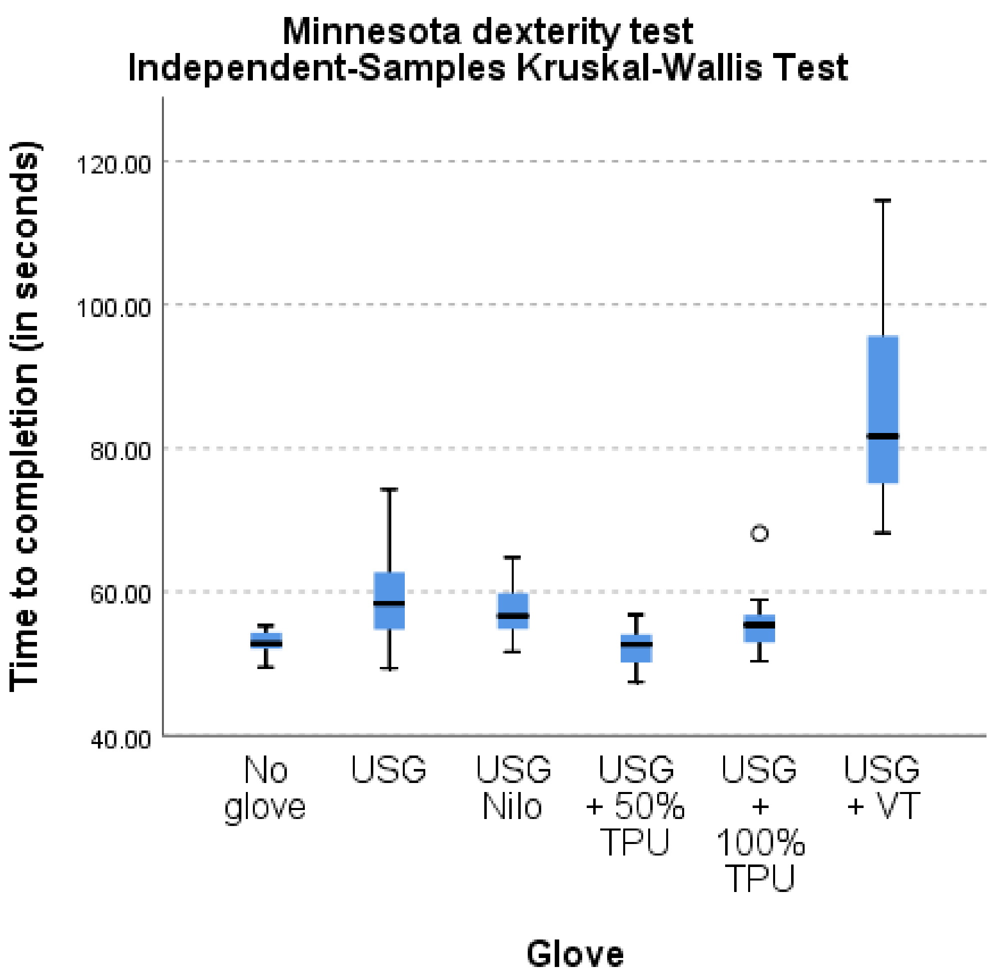


D.1.4 Minnesota dexterity test

Independent-Samples Kruskal-Wallis Test Summary

Total N	72
Test Statistic	42.799[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



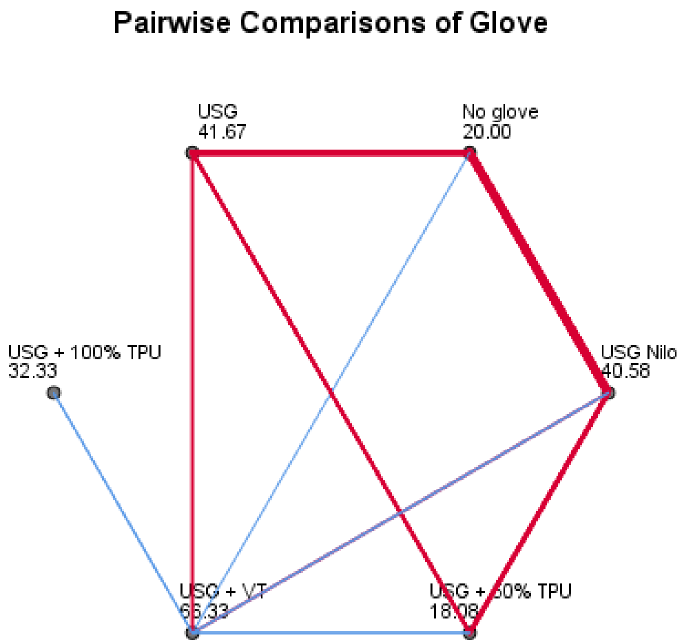
Pairwise Comparisons of Glove

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
USG + 50% TPU-No glove	1.917	8.544	.224	.823	1.000
USG + 50% TPU-USG + 100% TPU	-14.250	8.544	-1.668	.095	1.000
USG + 50% TPU-USG Nilo	22.500	8.544	2.633	.008	.127
USG + 50% TPU-USG	23.583	8.544	2.760	.006	.087
USG + 50% TPU-USG + VT	-48.250	8.544	-5.647	.000	.000
No glove-USG + 100% TPU	-12.333	8.544	-1.444	.149	1.000
No glove-USG Nilo	-20.583	8.544	-2.409	.016	.240
No glove-USG	-21.667	8.544	-2.536	.011	.168
No glove-USG + VT	-46.333	8.544	-5.423	.000	.000
USG + 100% TPU-USG Nilo	8.250	8.544	.966	.334	1.000
USG + 100% TPU-USG	9.333	8.544	1.092	.275	1.000
USG + 100% TPU-USG + VT	-34.000	8.544	-3.979	.000	.001
USG Nilo-USG	1.083	8.544	.127	.899	1.000
USG Nilo-USG + VT	-25.750	8.544	-3.014	.003	.039
USG-USG + VT	-24.667	8.544	-2.887	.004	.058

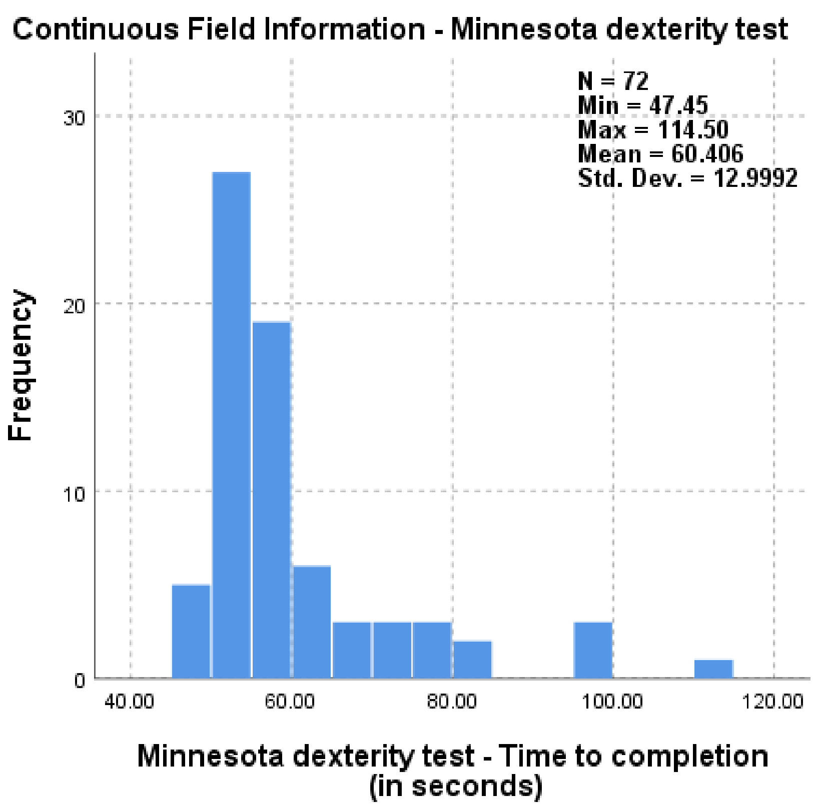
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



Each node shows the sample average rank of Glove.



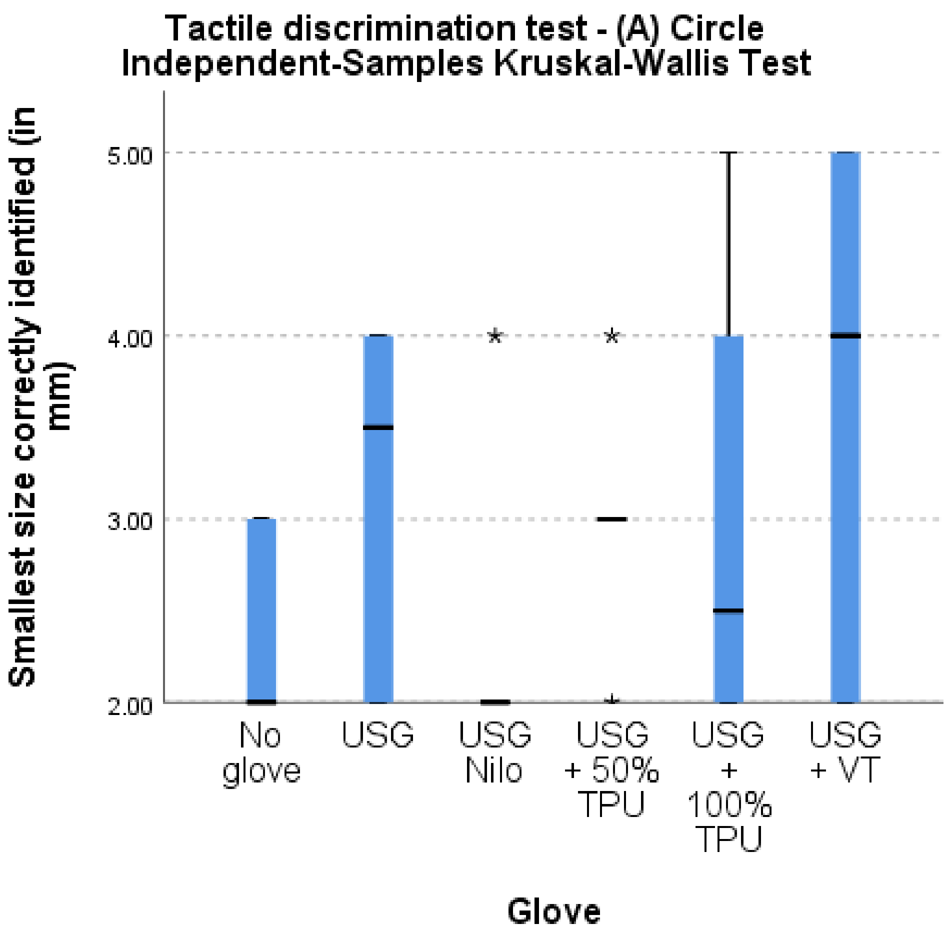
D.1.5 *Tactile discrimination test - (A) Circle*

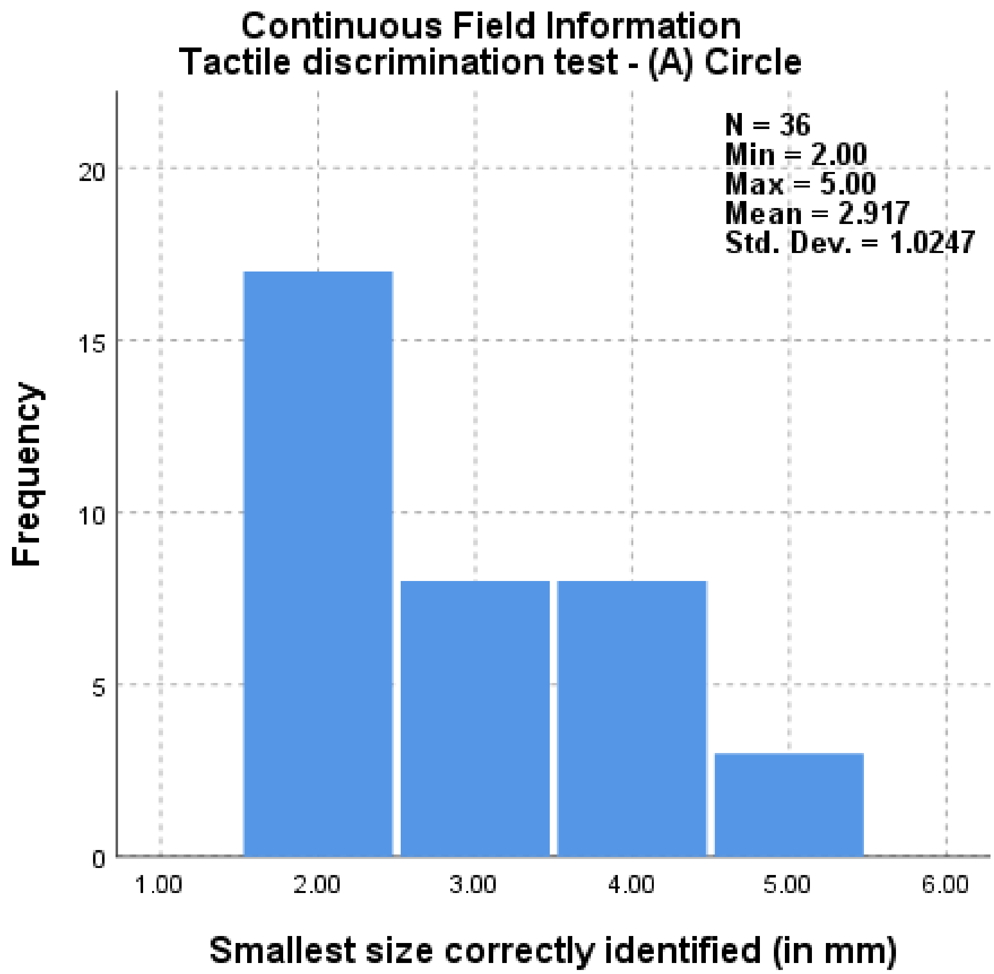
Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	7.100[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.213

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.




Continuous Field Information
Tactile discrimination test - (A) Circle
N = 36
Min = 2.00
Max = 5.00
Mean = 2.917
Std. Dev. = 1.0247
Frequency
20
15
10
5
0
1.00
2.00
3.00
4.00
5.00
6.00
Smallest size correctly identified (in mm)

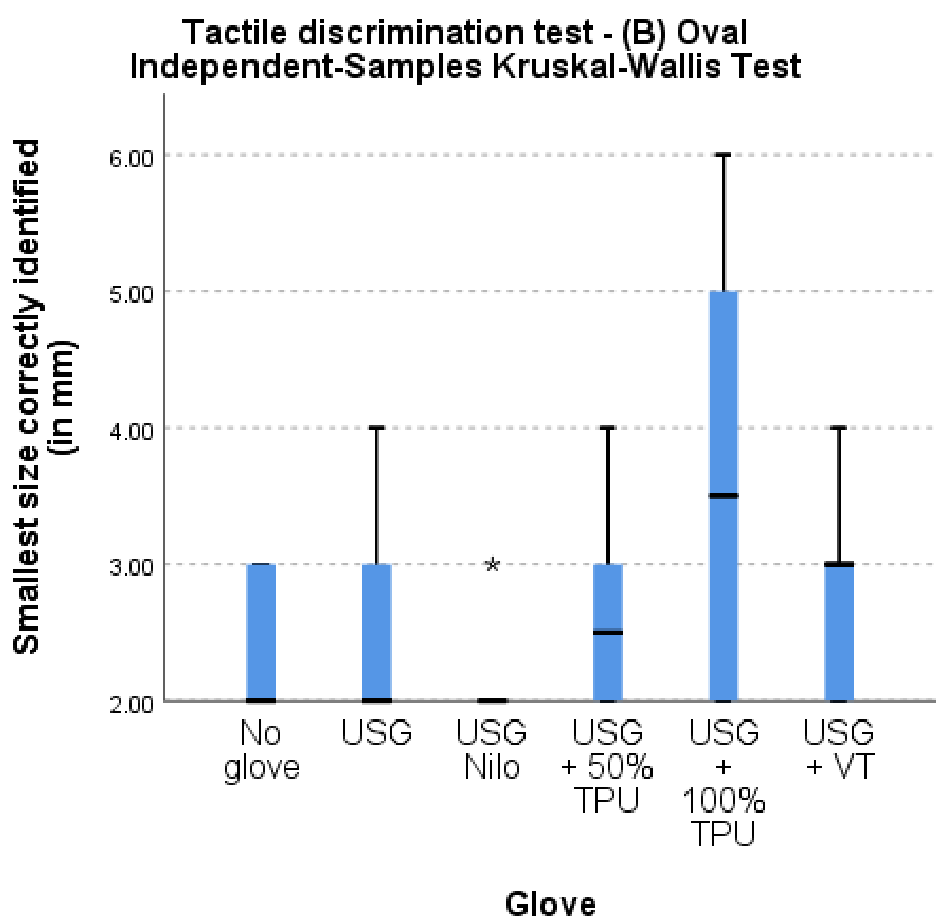
D.1.6 Tactile discrimination test - (B) Oval

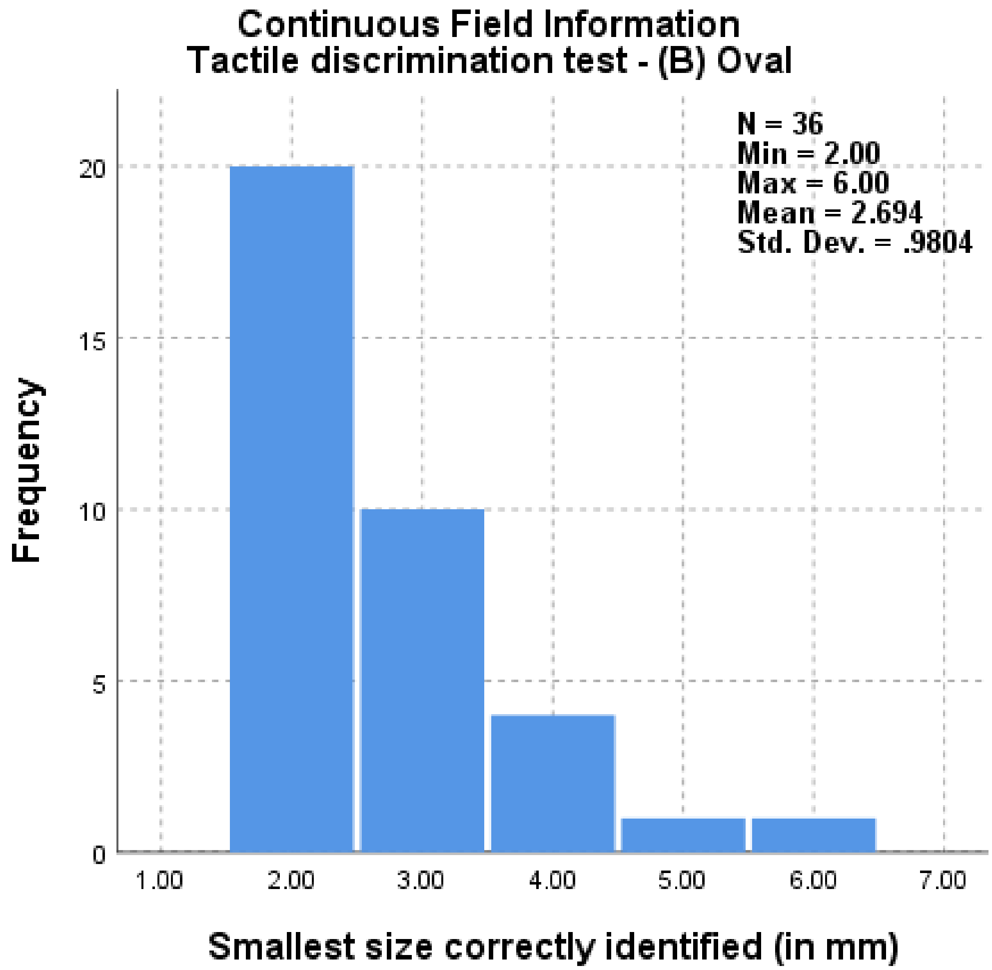
Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	6.391[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.270

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





Continuous Field Information
Tactile discrimination test - (B) Oval
N = 36
Min = 2.00
Max = 6.00
Mean = 2.694
Std. Dev. = .9804
Frequency
0
5
10
15
20
1.00
2.00
3.00
4.00
5.00
6.00
7.00
Smallest size correctly identified (in mm)

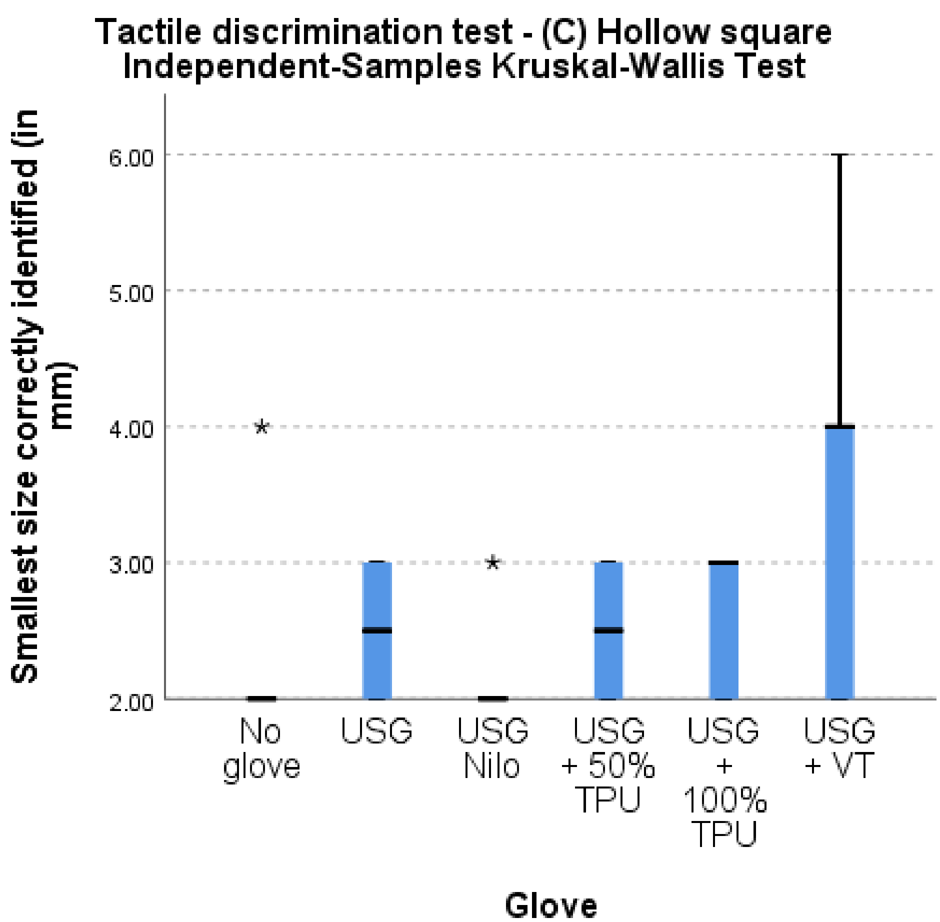
D.1.7 Tactile discrimination test - (C) Hollow square

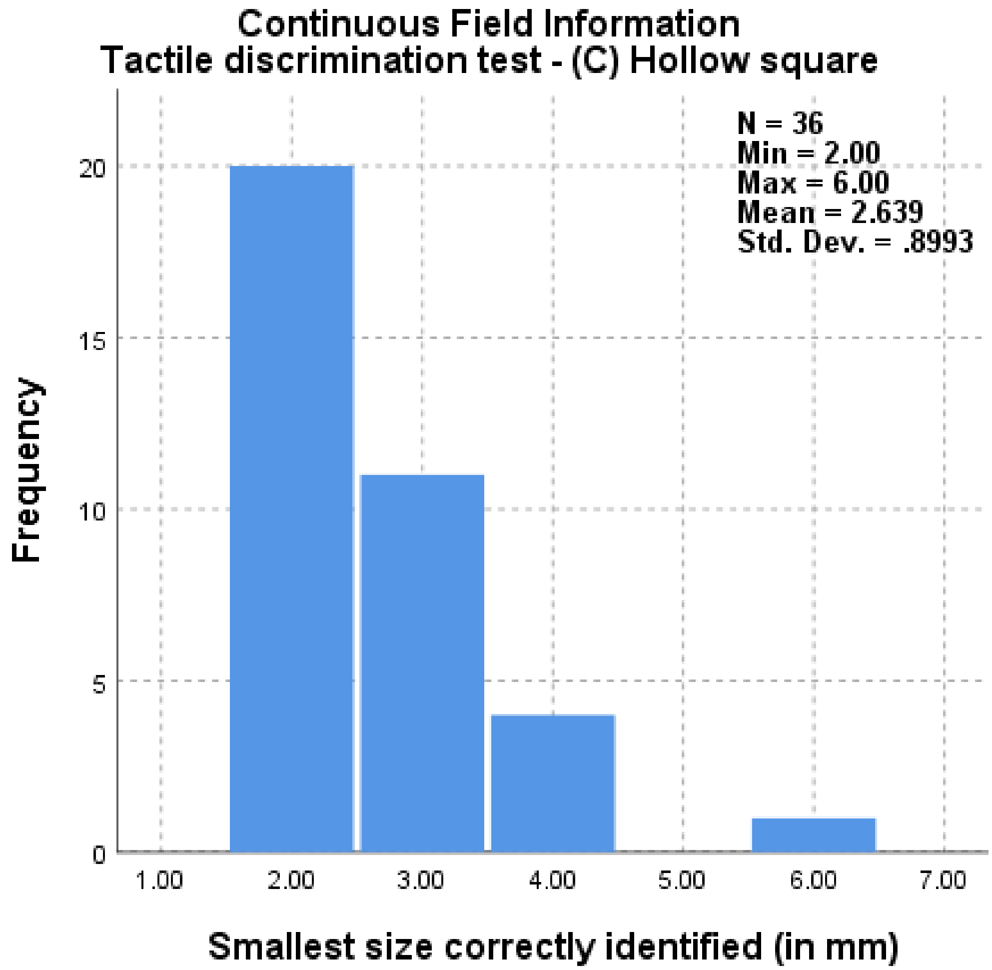
Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	7.597[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.180

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.




Continuous Field Information
Tactile discrimination test - (C) Hollow square
N = 36
Min = 2.00
Max = 6.00
Mean = 2.639
Std. Dev. = .8993
Frequency
0
5
10
15
20
1.00
2.00
3.00
4.00
5.00
6.00
7.00
Smallest size correctly identified (in mm)

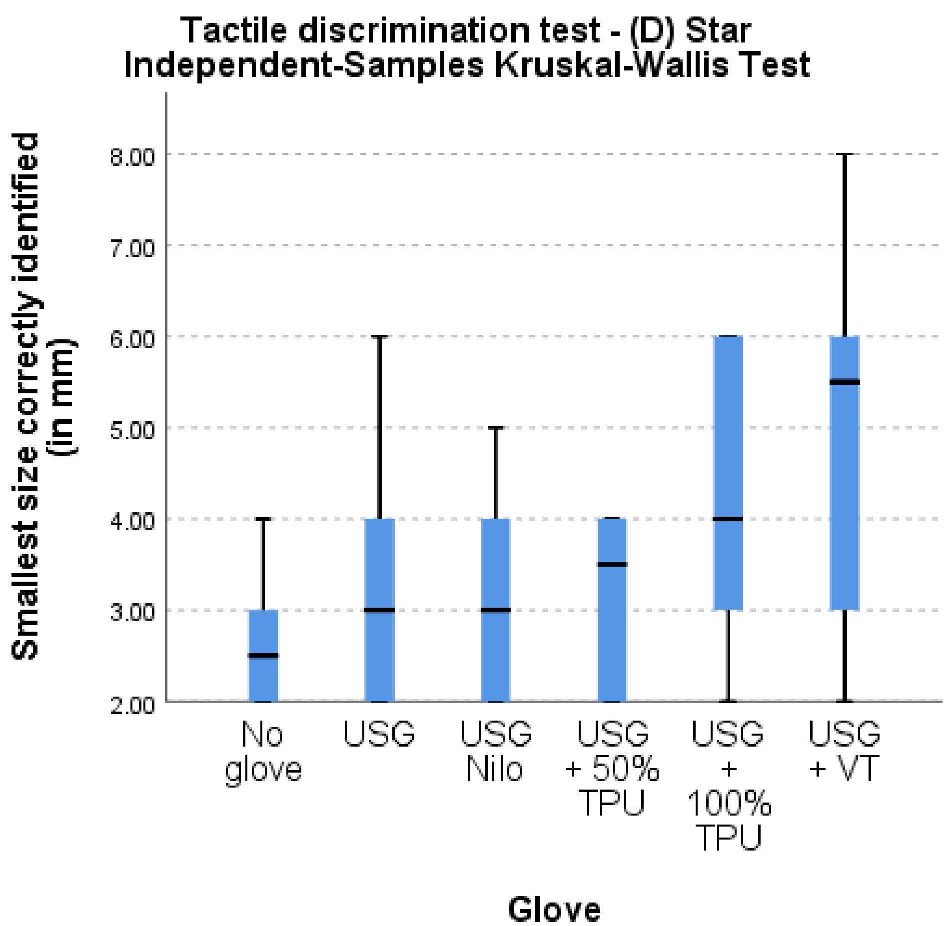
D.1.8 Tactile discrimination test - (D) Star

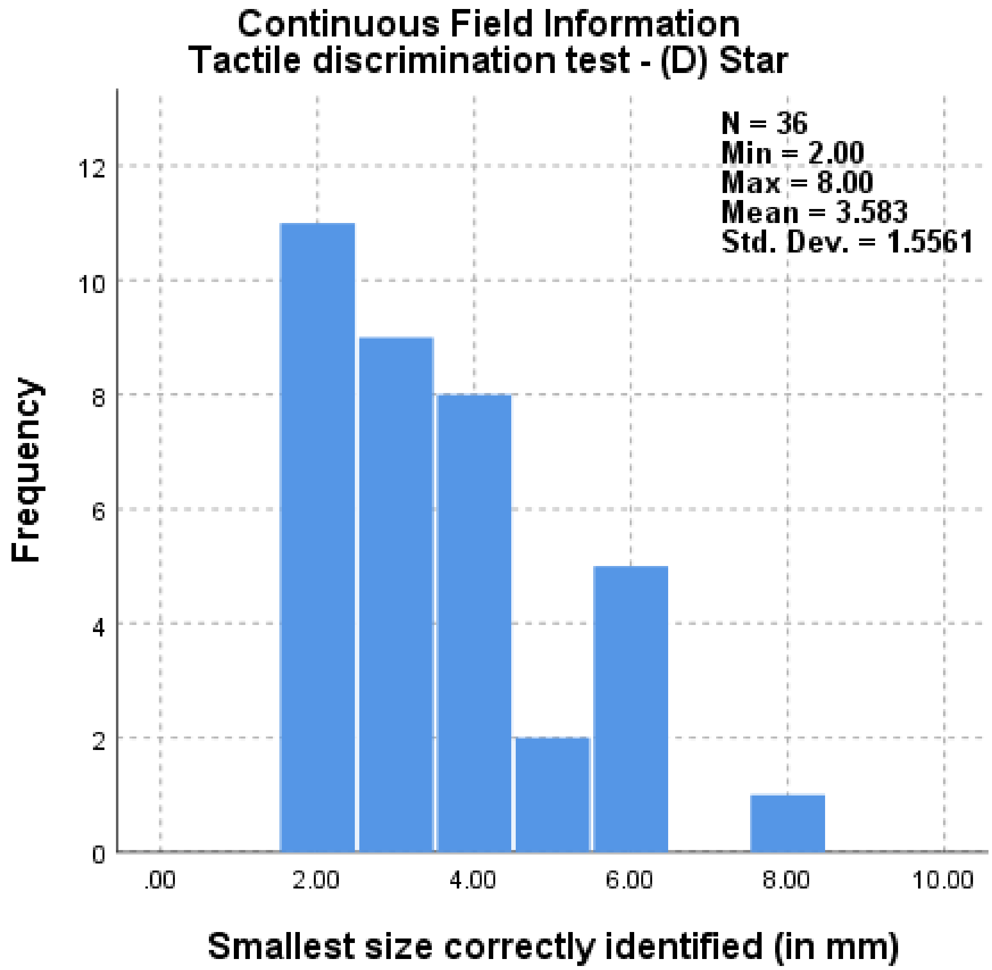
Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	6.632[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.249

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





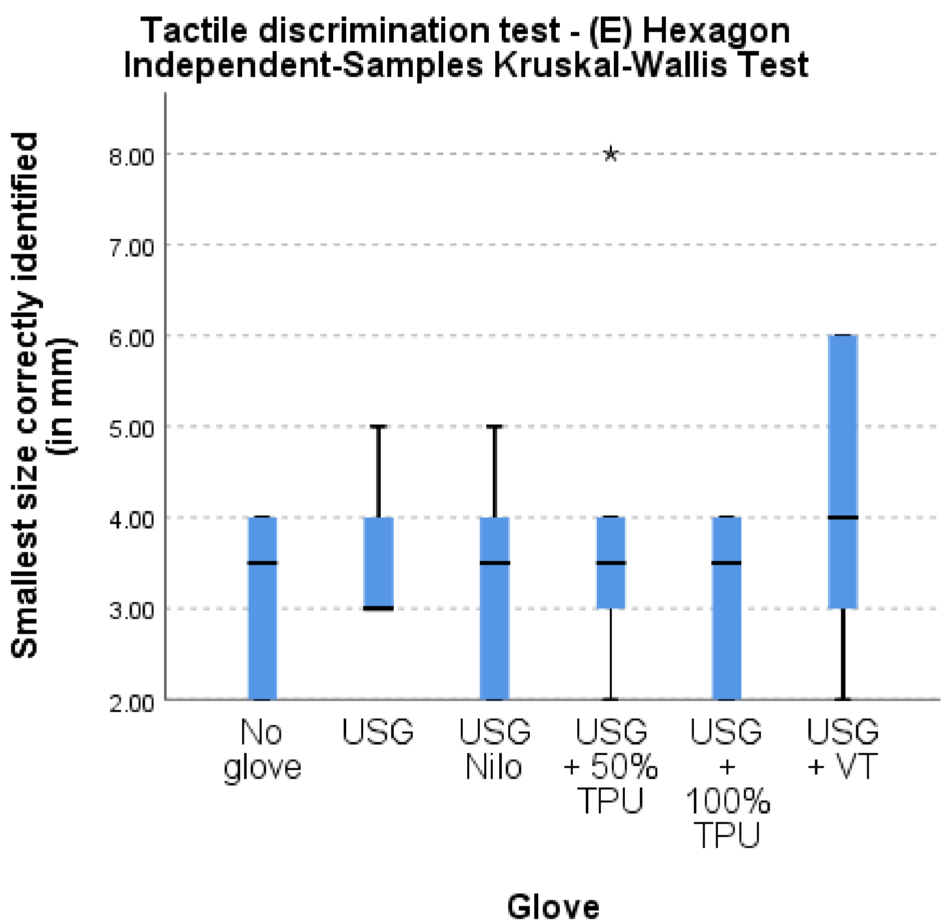
Continuous Field Information
Tactile discrimination test - (D) Star
N = 36
Min = 2.00
Max = 8.00
Mean = 3.583
Std. Dev. = 1.5561
Frequency
12
10
8
6
4
2
0
.00
2.00
4.00
6.00
8.00
10.00
Smallest size correctly identified (in mm)

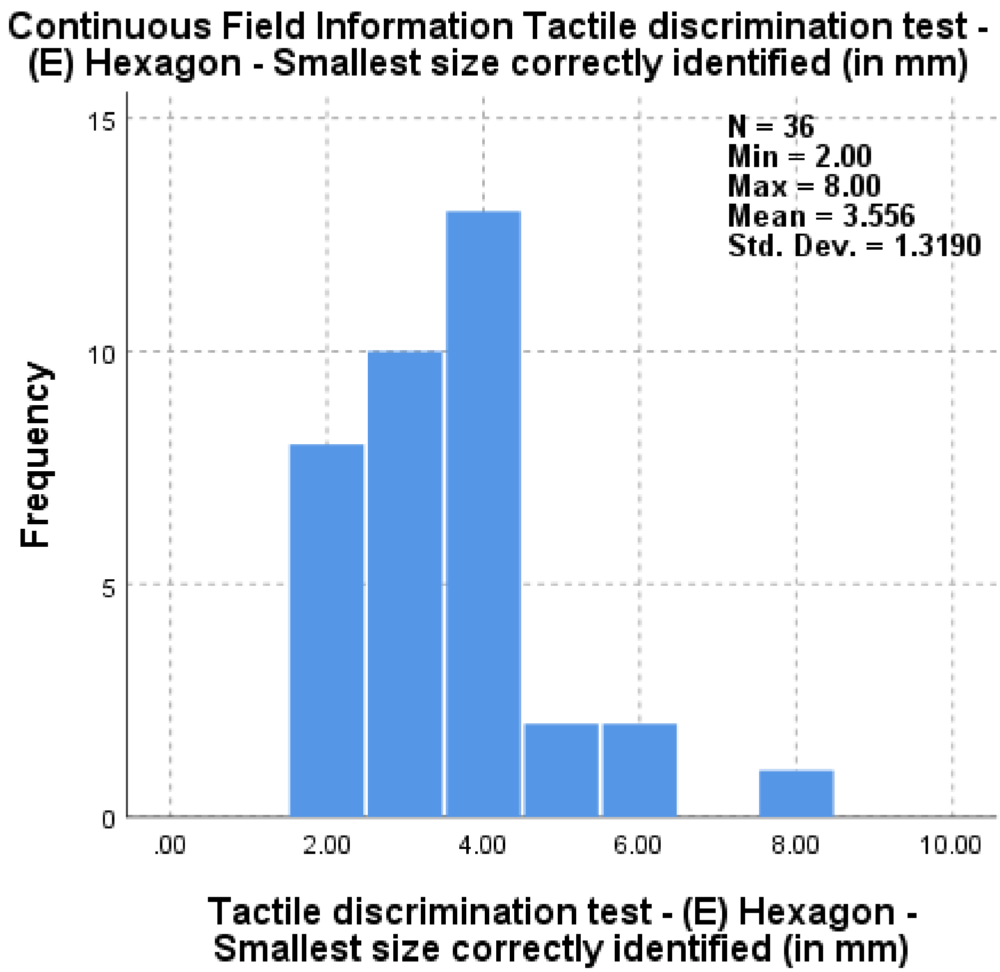
D.1.9 Tactile discrimination test - (E) Hexagon

Independent-Samples Kruskal-Wallis Test Summary	
Total N	36
Test Statistic	1.825[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.873

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.



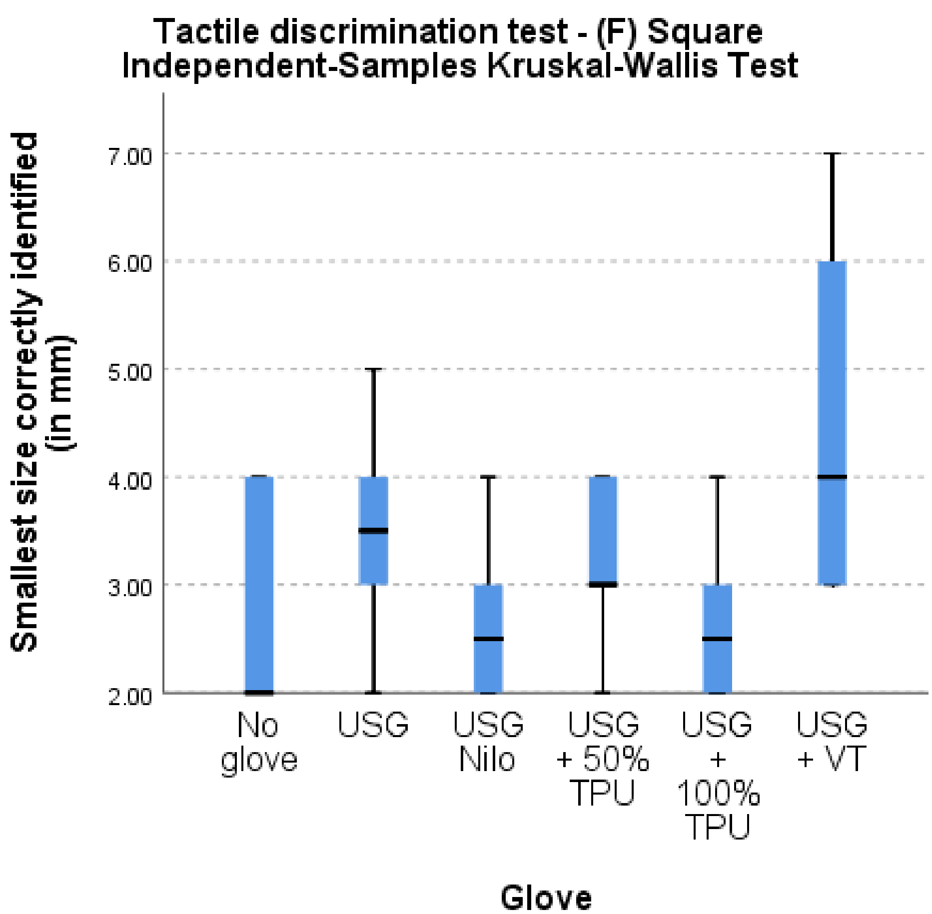

Continuous Field Information Tactile discrimination test - (E) Hexagon - Smallest size correctly identified (in mm)
N = 36
Min = 2.00
Max = 8.00
Mean = 3.556
Std. Dev. = 1.3190
Frequency
15
10
5
0
.00
2.00
4.00
6.00
8.00
10.00
Tactile discrimination test - (E) Hexagon - Smallest size correctly identified (in mm)

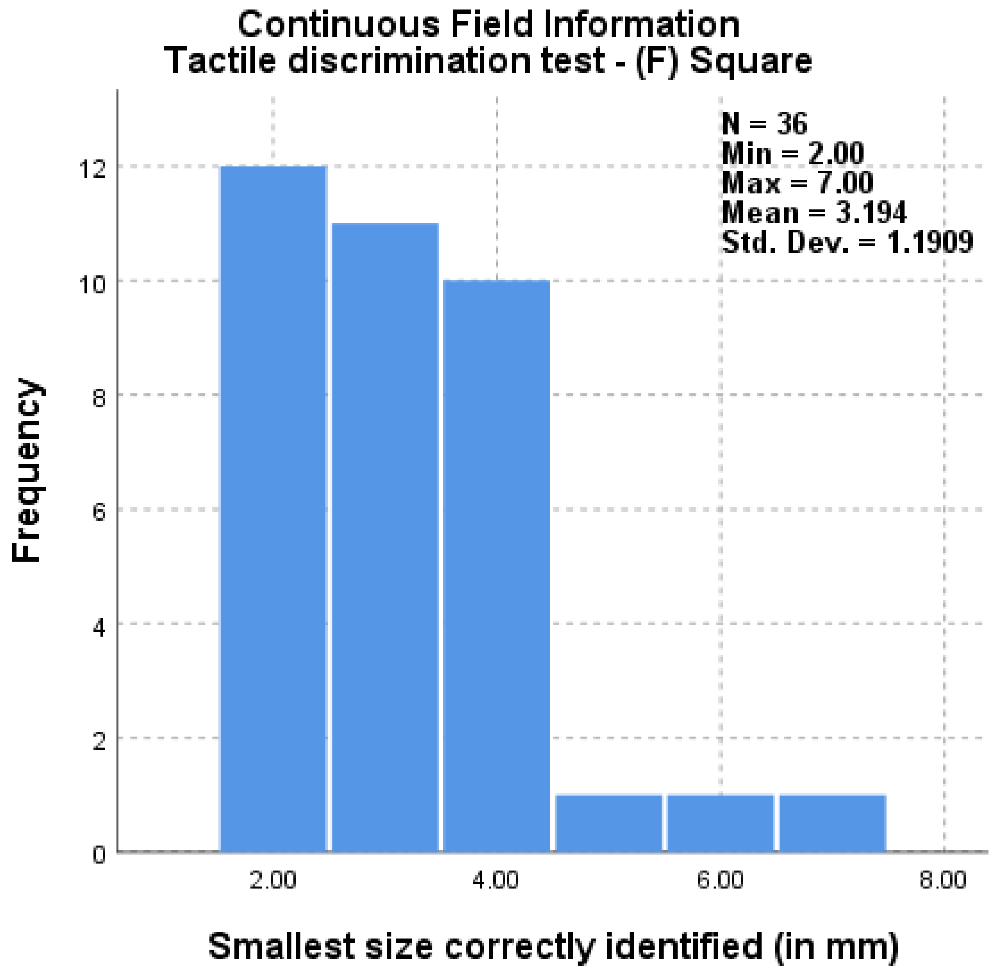
D.1.10 Tactile discrimination test - (F) Square

Independent-Samples Kruskal-Wallis Test Summary	
Total N	36
Test Statistic	9.202[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.101

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





Continuous Field Information
Tactile discrimination test - (F) Square
N = 36
Min = 2.00
Max = 7.00
Mean = 3.194
Std. Dev. = 1.1909
Frequency
0
2
4
6
8
10
12
2.00
4.00
6.00
8.00
Smallest size correctly identified (in mm)

D.2 Error rate data

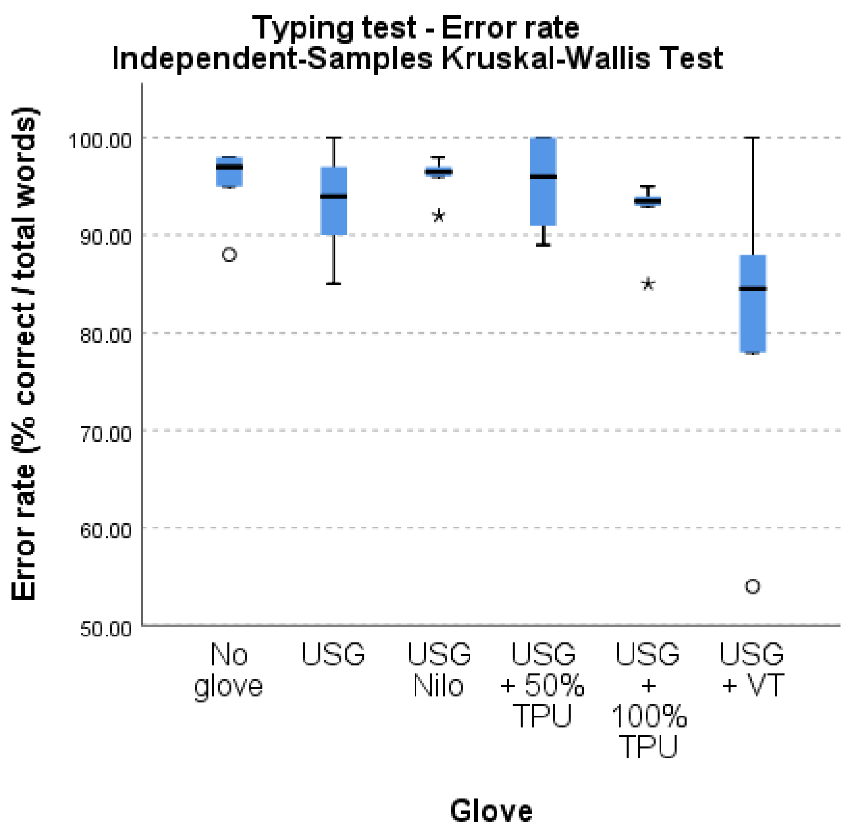
D.2.1 Typing test - Error rate

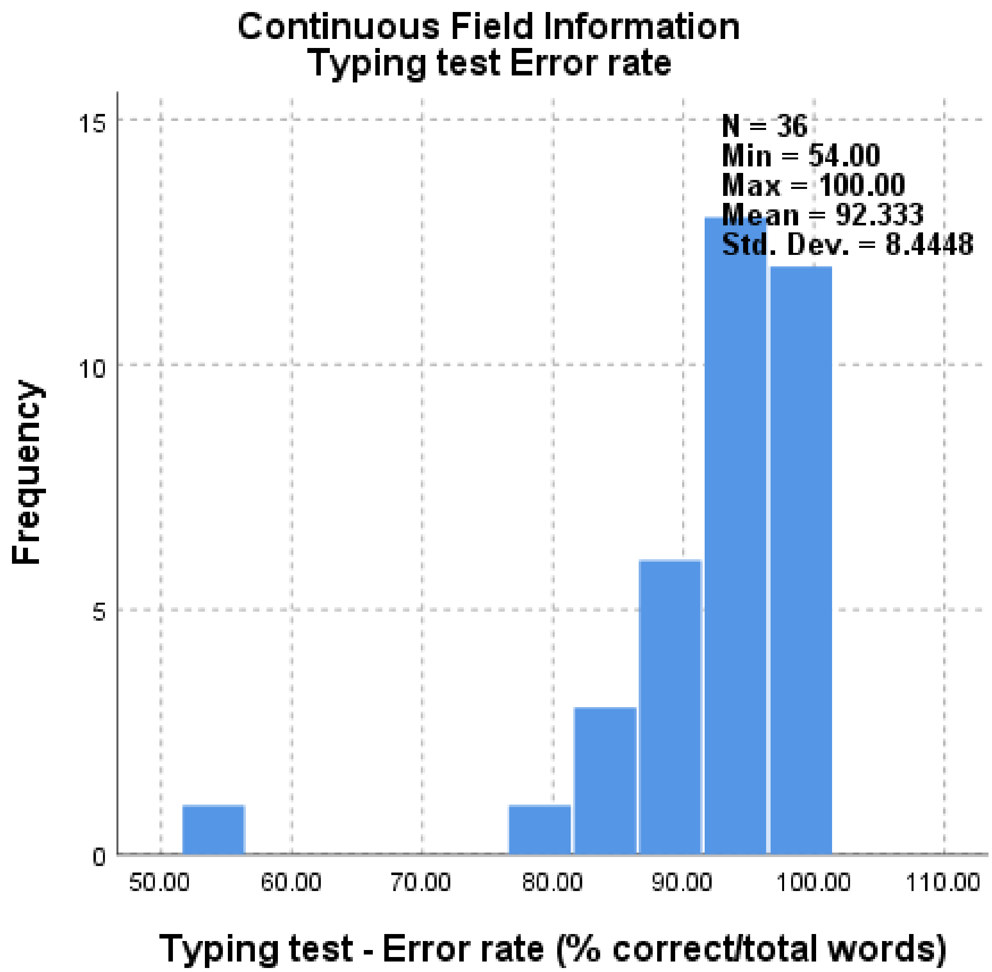
Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	9.735[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.083

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

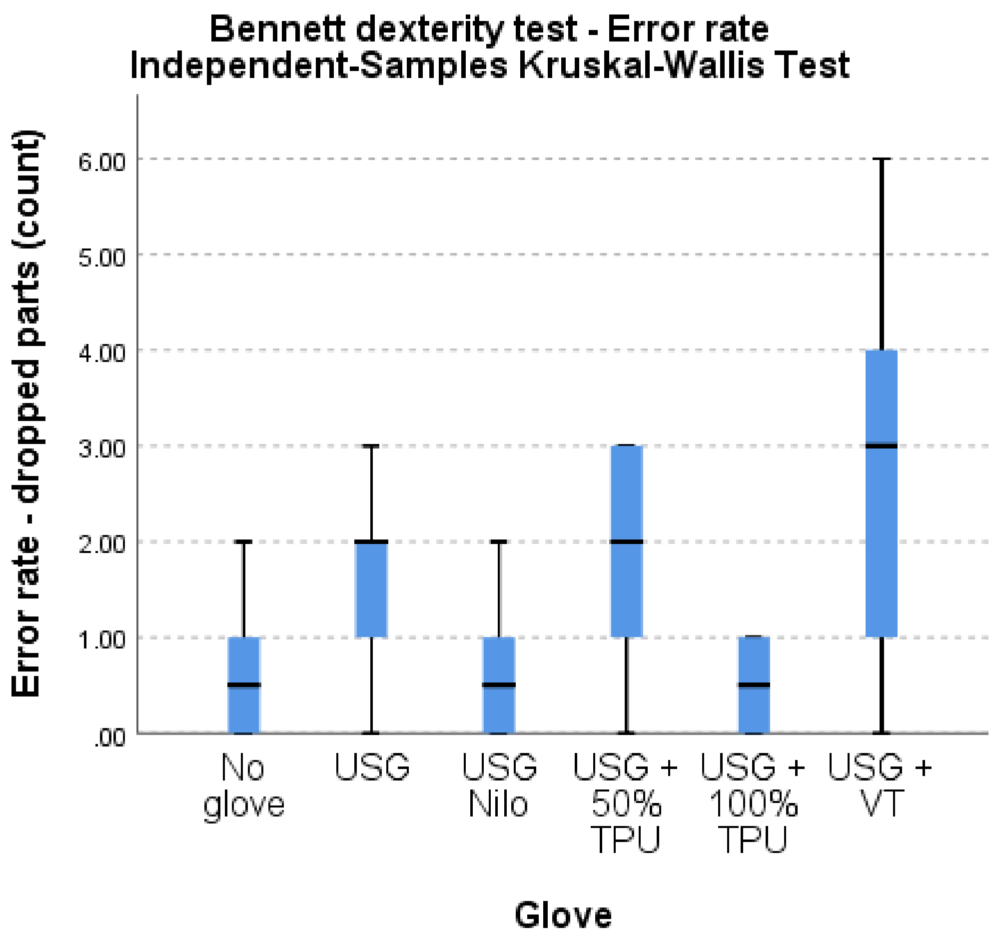



Continuous Field Information
Typing test Error rate
N = 36
Min = 54.00
Max = 100.00
Mean = 92.333
Std. Dev. = 8.4448
Frequency
15
10
5
0
50.00
60.00
70.00
80.00
90.00
100.00
110.00
Typing test - Error rate (% correct/total words)

D.2.2 Bennett dexterity test - Error rate

Independent-Samples Kruskal-Wallis Test Summary	
Total N	36
Test Statistic	11.365[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.045

a. The test statistic is adjusted for ties.



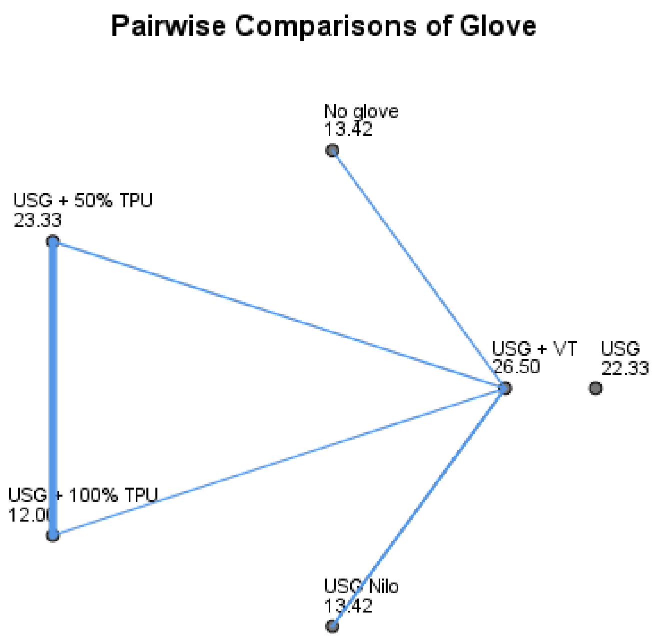
Pairwise Comparisons of Glove

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
USG + 100% TPU-No glove	1.417	5.873	.241	.809	1.000
USG + 100% TPU-USG Nilo	1.417	5.873	.241	.809	1.000
USG + 100% TPU-USG	10.333	5.873	1.760	.078	1.000
USG + 100% TPU-USG + 50% TPU	11.333	5.873	1.930	.054	.805
USG + 100% TPU-USG + VT	-14.500	5.873	-2.469	.014	.203
No glove-USG Nilo	.000	5.873	.000	1.000	1.000
No glove-USG	-8.917	5.873	-1.518	.129	1.000
No glove-USG + 50% TPU	-9.917	5.873	-1.689	.091	1.000
No glove-USG + VT	-13.083	5.873	-2.228	.026	.388
USG Nilo-USG	8.917	5.873	1.518	.129	1.000
USG Nilo-USG + 50% TPU	-9.917	5.873	-1.689	.091	1.000
USG Nilo-USG + VT	-13.083	5.873	-2.228	.026	.388
USG-USG + 50% TPU	-1.000	5.873	-.170	.865	1.000
USG-USG + VT	-4.167	5.873	-.709	.478	1.000
USG + 50% TPU-USG + VT	-3.167	5.873	-.539	.590	1.000

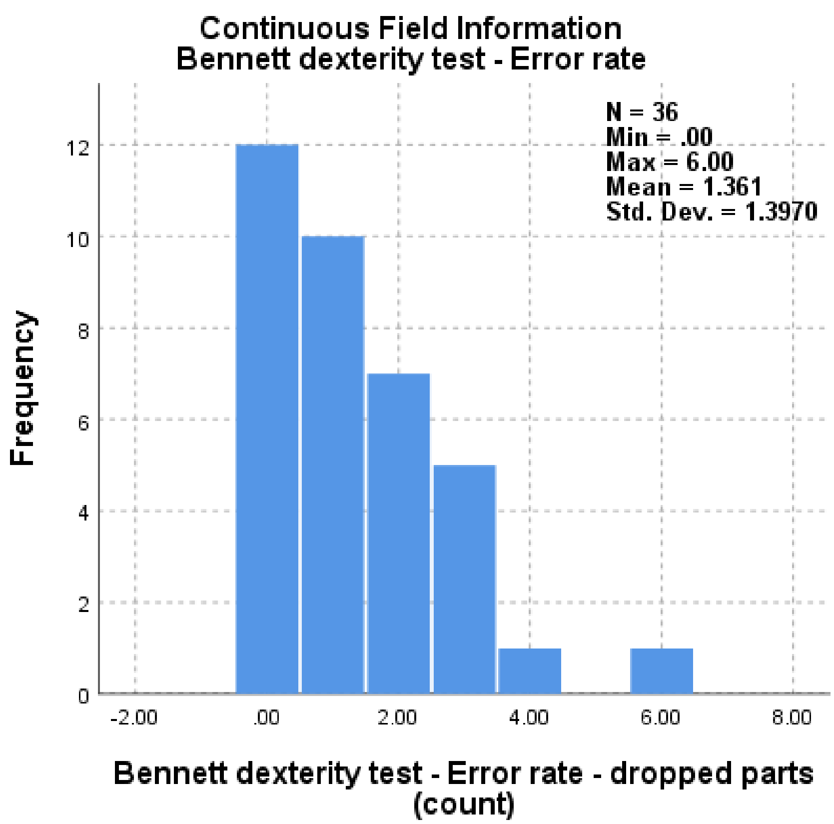
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



Each node shows the sample average rank of Glove.



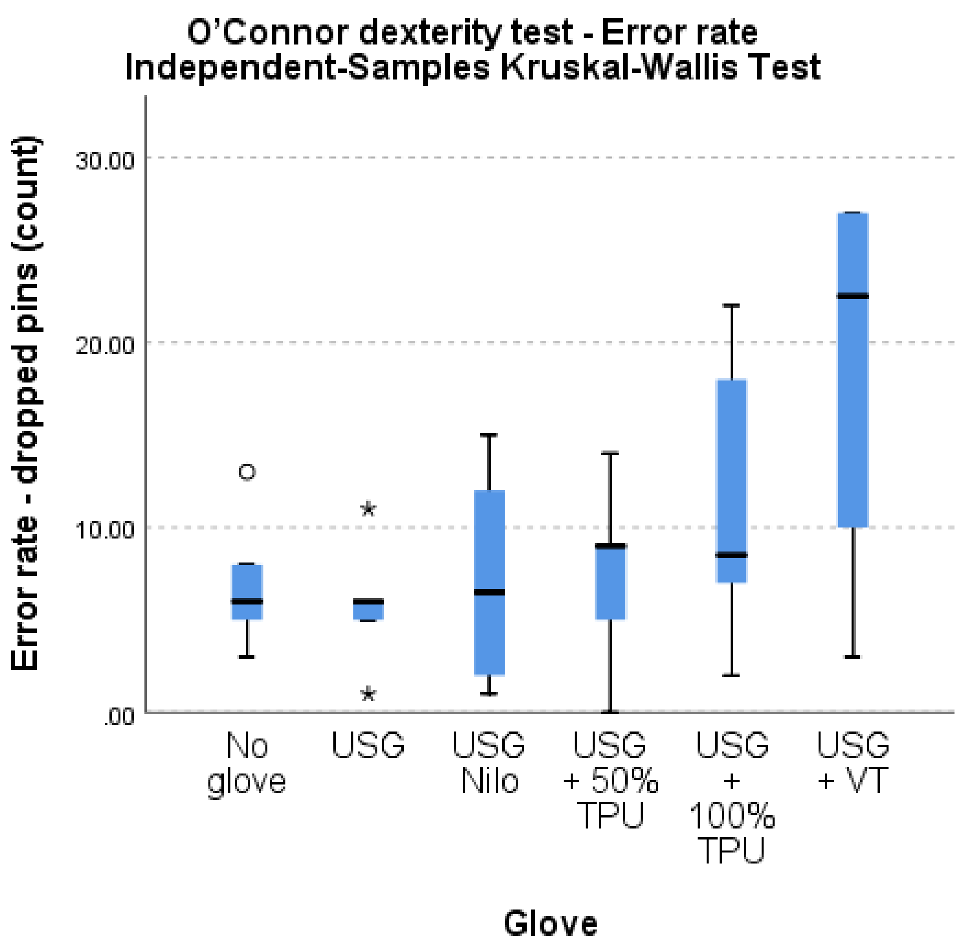
D.2.3 O'Connor dexterity test - Error rate

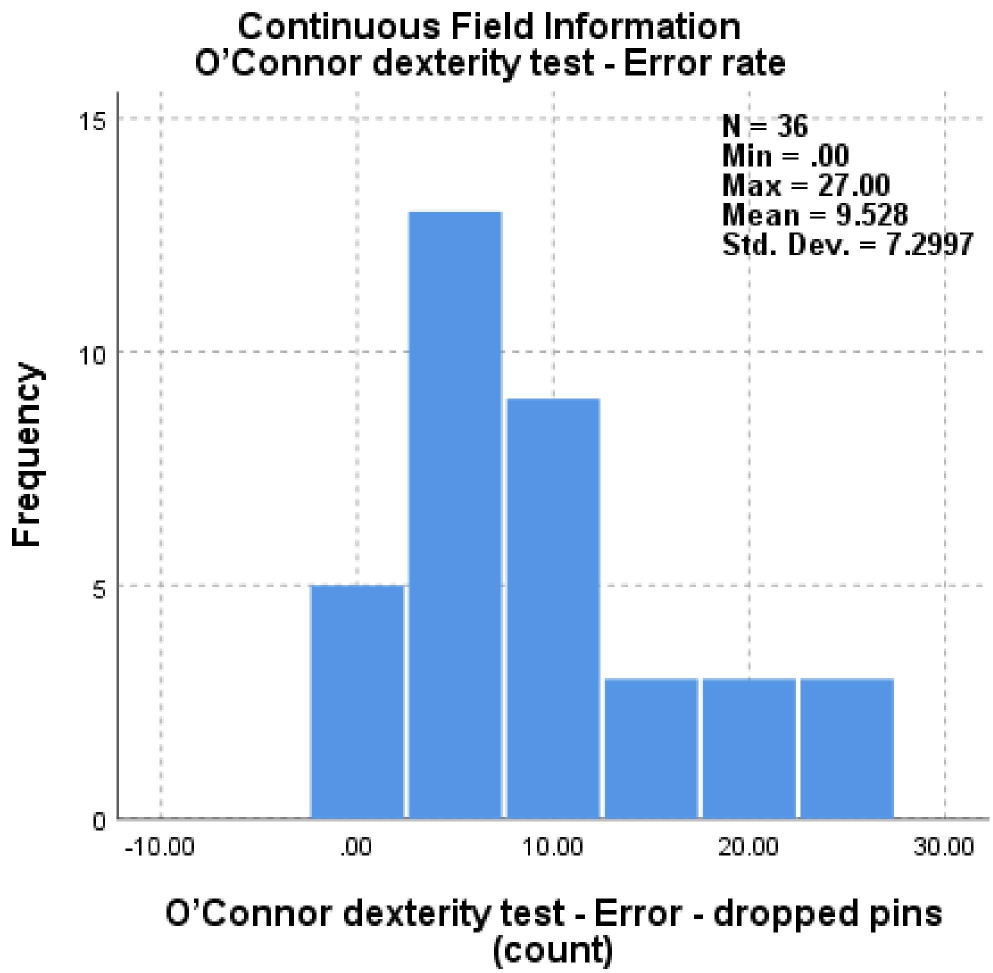
Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	7.698[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.174

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.




Continuous Field Information
O'Connor dexterity test - Error rate
N = 36
Min = .00
Max = 27.00
Mean = 9.528
Std. Dev. = 7.2997
Frequency
15
10
5
0
-10.00
.00
10.00
20.00
30.00
O'Connor dexterity test - Error - dropped pins (count)

D.3 Task NASA-TLX data

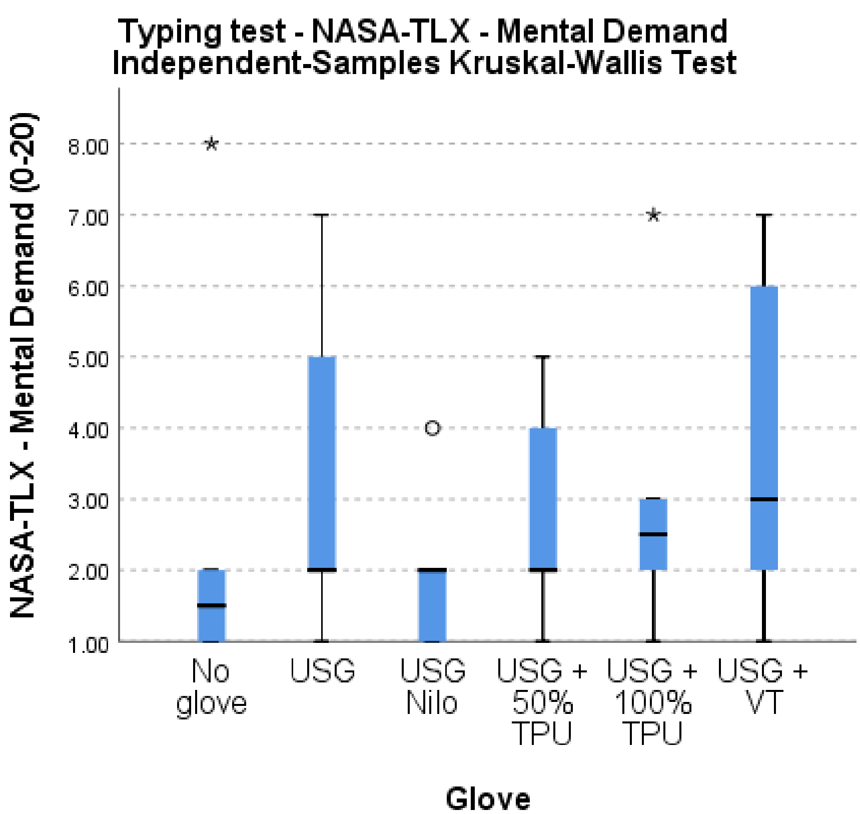
D.3.1 Typing test - NASA-TLX - Mental Demand

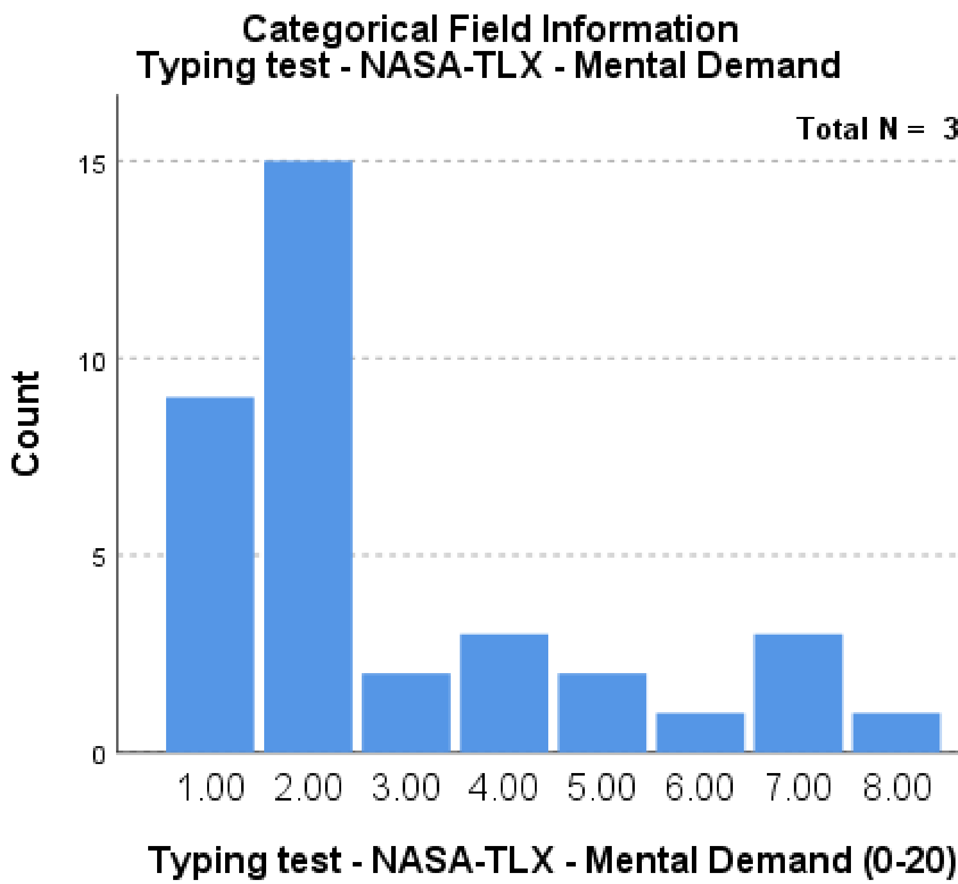
Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	3.156[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.676

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





Typing test - NASA-TLX - Mental Demand (0-20) field is ordinal but is treated as continuous in the test.

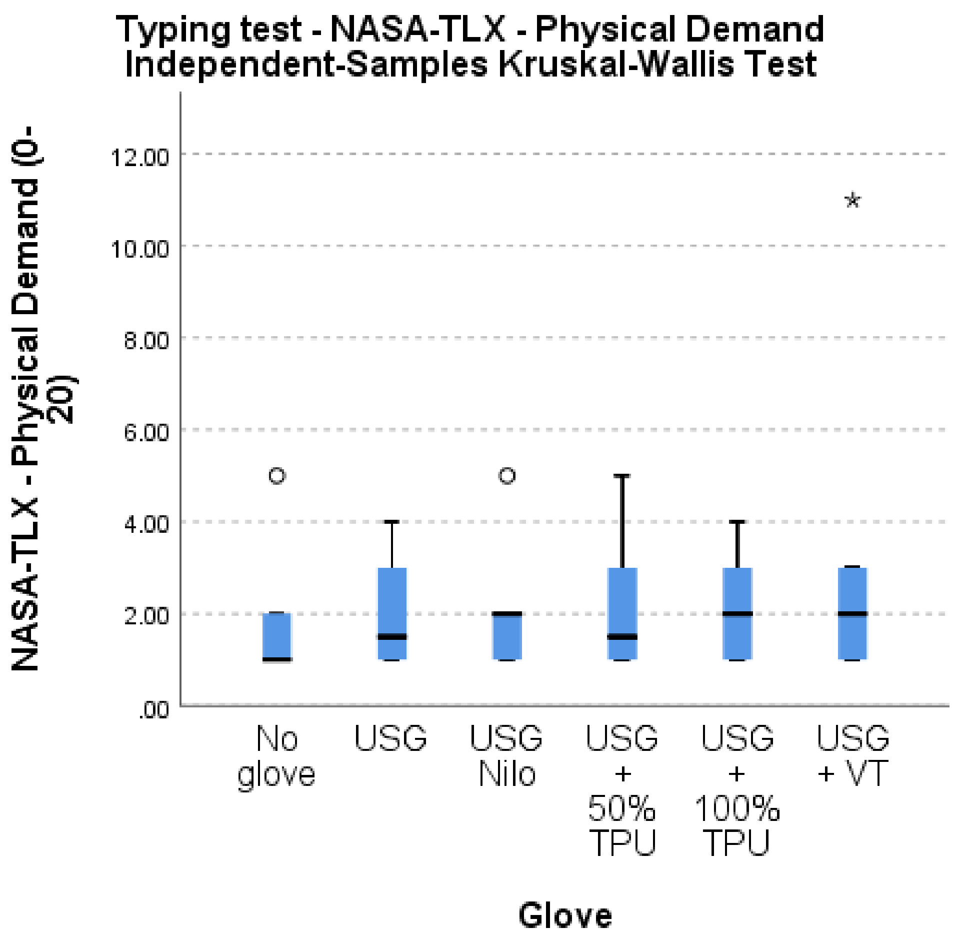
D.3.2 Typing test - NASA-TLX - Physical Demand

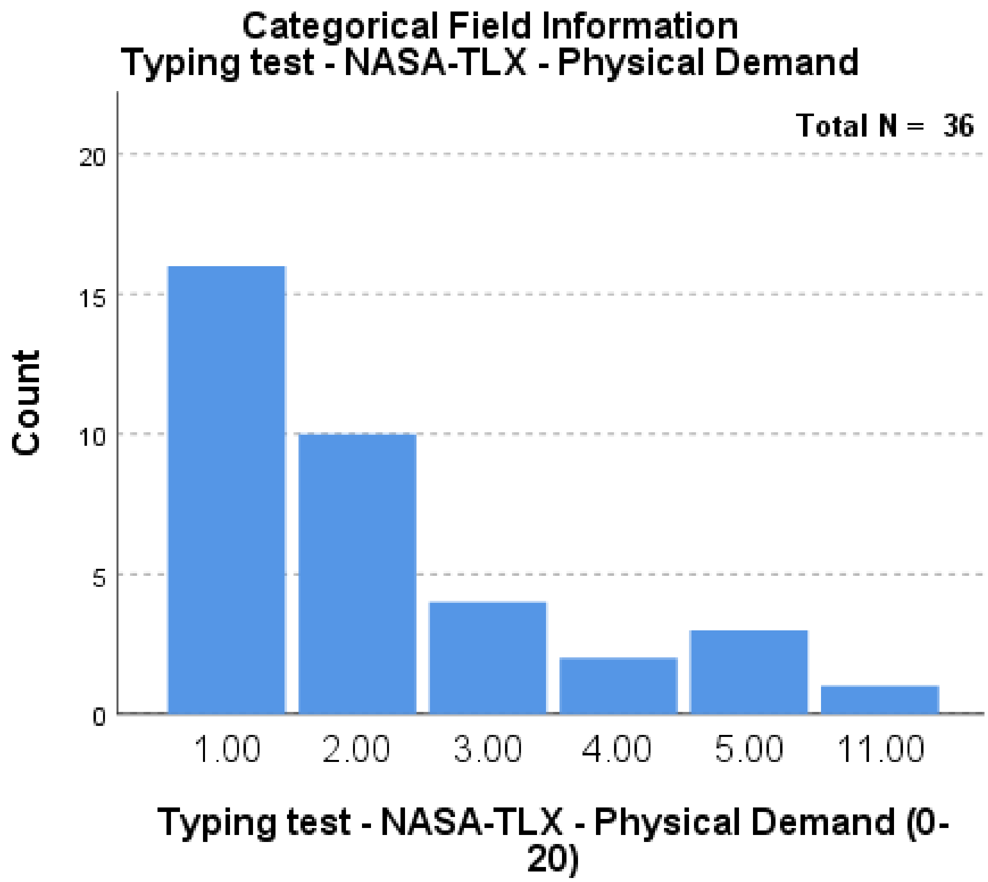
Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	1.290[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.936

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





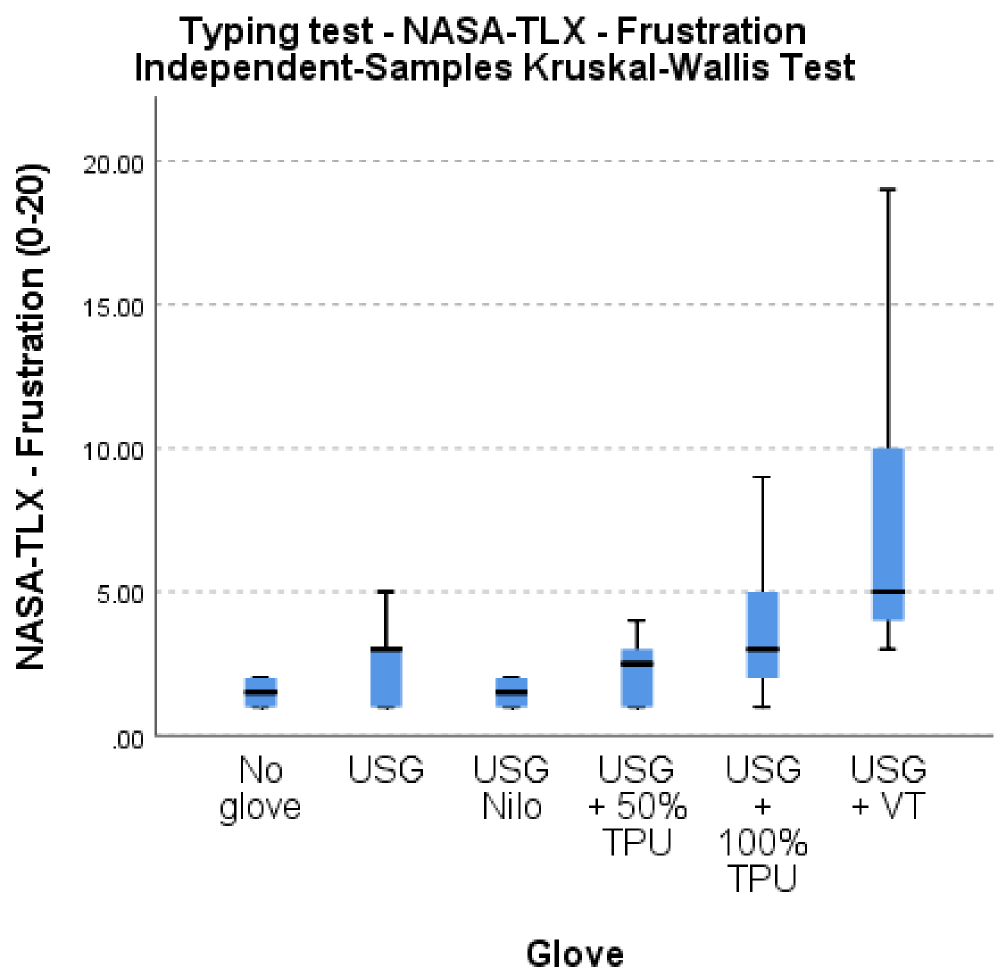
Typing test - NASA-TLX - Physical Demand (0-20) field is ordinal but is treated as continuous in the test.

D.3.3 Typing test - NASA-TLX - Frustration

Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	15.954[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.007

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

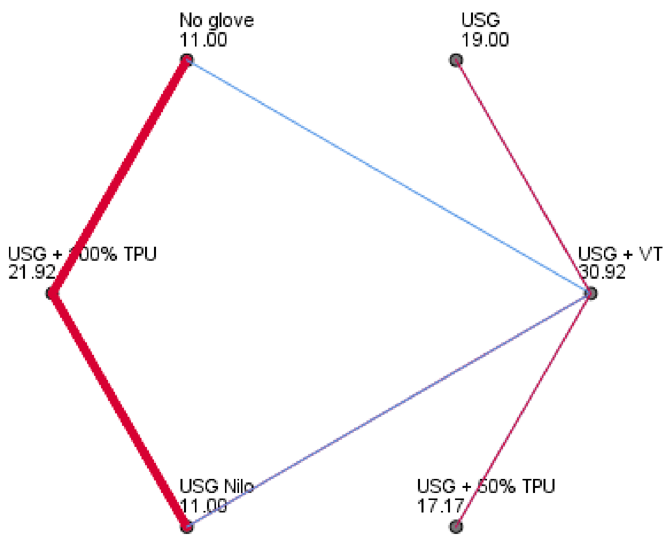
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
No glove-USG Nilo	.000	5.929	.000	1.000	1.000
No glove-USG + 50% TPU	-6.167	5.929	-1.040	.298	1.000
No glove-USG	-8.000	5.929	-1.349	.177	1.000
No glove-USG + 100% TPU	-10.917	5.929	-1.841	.066	.984
No glove-USG + VT	-19.917	5.929	-3.359	.001	.012
USG Nilo-USG + 50% TPU	-6.167	5.929	-1.040	.298	1.000
USG Nilo-USG	8.000	5.929	1.349	.177	1.000
USG Nilo-USG + 100% TPU	-10.917	5.929	-1.841	.066	.984
USG Nilo-USG + VT	-19.917	5.929	-3.359	.001	.012
USG + 50% TPU-USG	1.833	5.929	.309	.757	1.000
USG + 50% TPU-USG + 100% TPU	-4.750	5.929	-.801	.423	1.000
USG + 50% TPU-USG + VT	-13.750	5.929	-2.319	.020	.306
USG-USG + 100% TPU	-2.917	5.929	-.492	.623	1.000
USG-USG + VT	-11.917	5.929	-2.010	.044	.666
USG + 100% TPU-USG + VT	-9.000	5.929	-1.518	.129	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

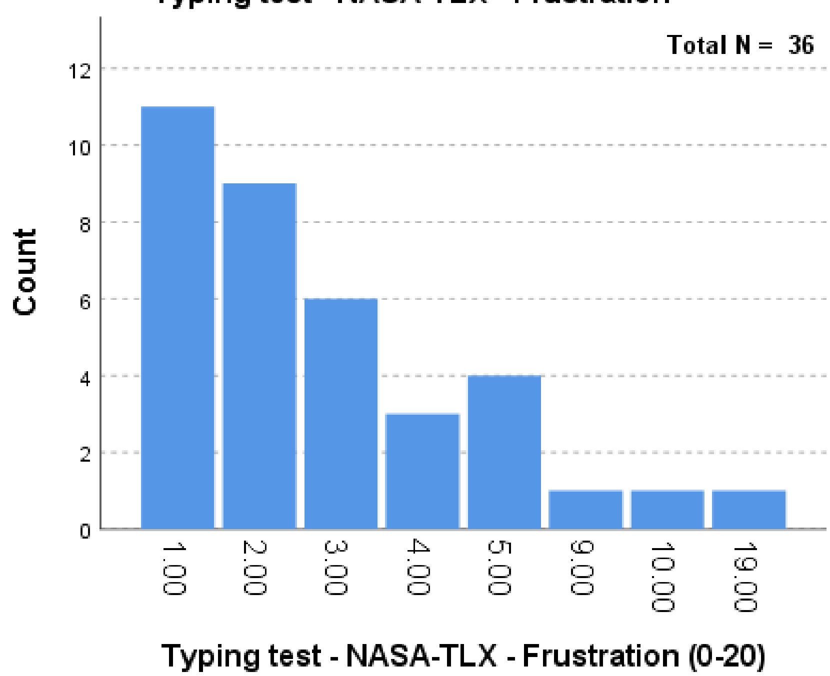
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.

Categorical Field Information
Typing test - NASA-TLX - Frustration



Typing test - NASA-TLX - Frustration (0-20) field is ordinal but is treated as continuous in the test.

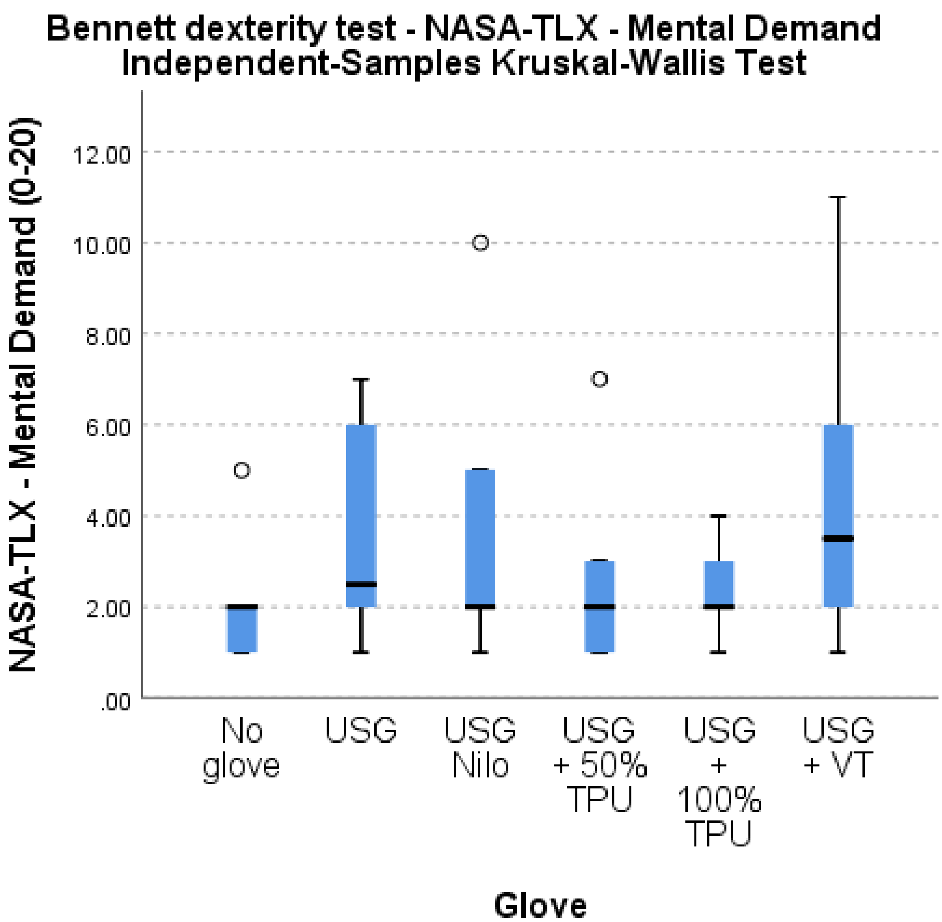
D.3.4 Bennett dexterity test - NASA-TLX - Mental Demand

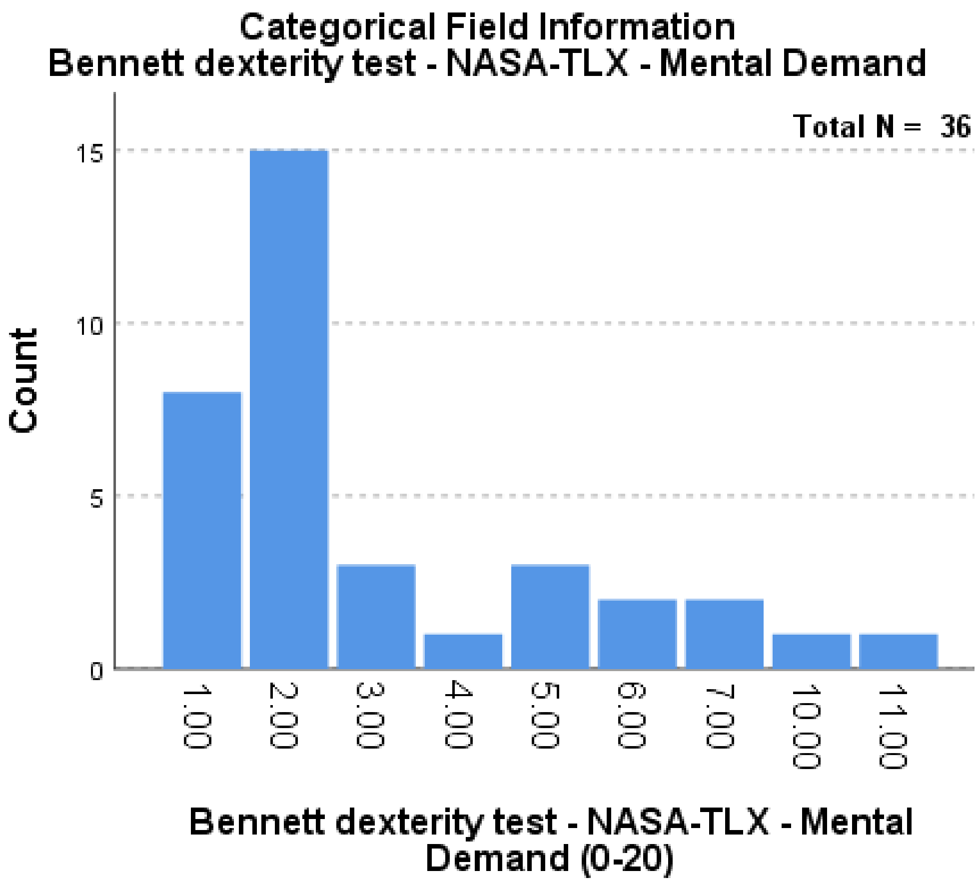
Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	2.548[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.769

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





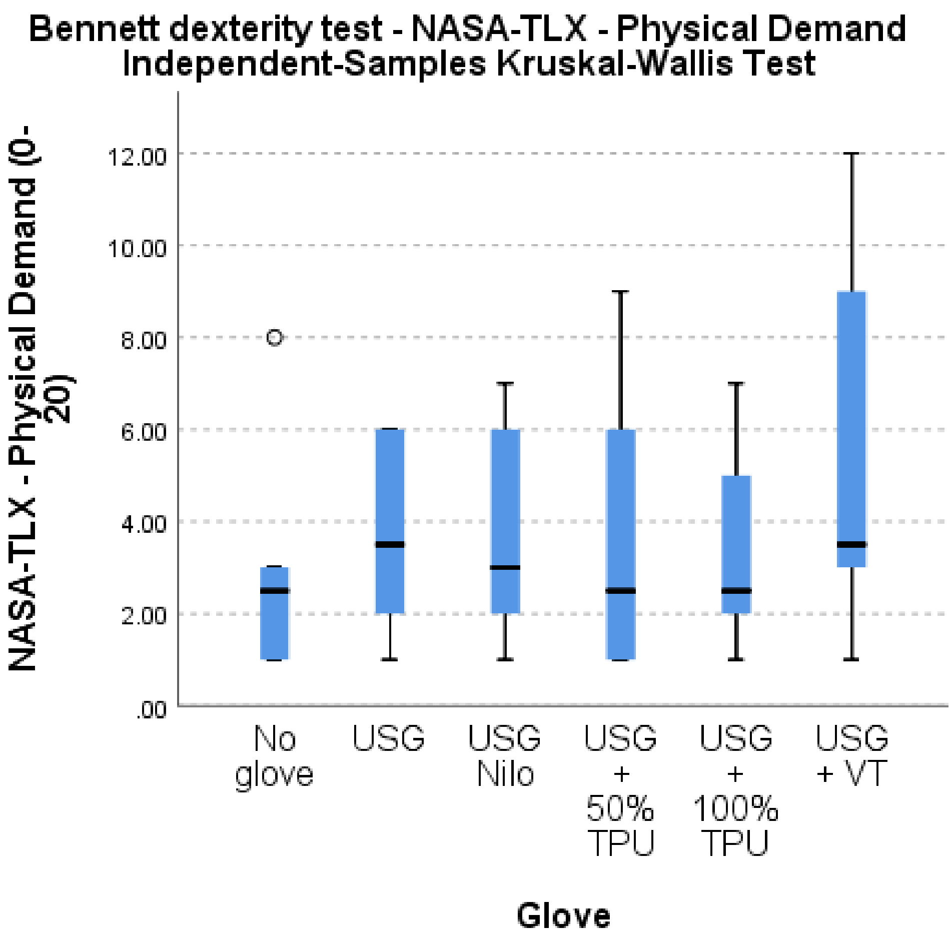
Bennett dexterity test - NASA-TLX - Mental Demand (0-20) field is ordinal but is treated as continuous in the test.

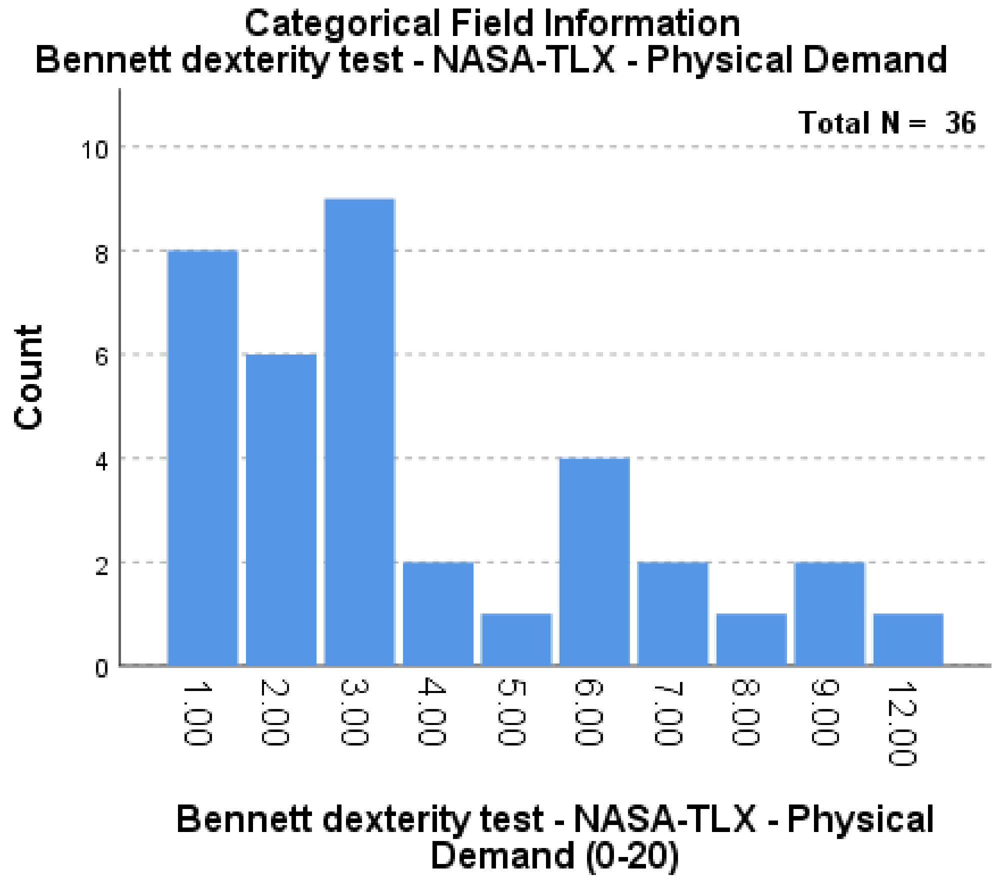
D.3.5 Bennett dexterity test - NASA-TLX - Physical Demand

Independent-Samples Kruskal-Wallis Test Summary	
Total N	36
Test Statistic	1.929[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.859

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





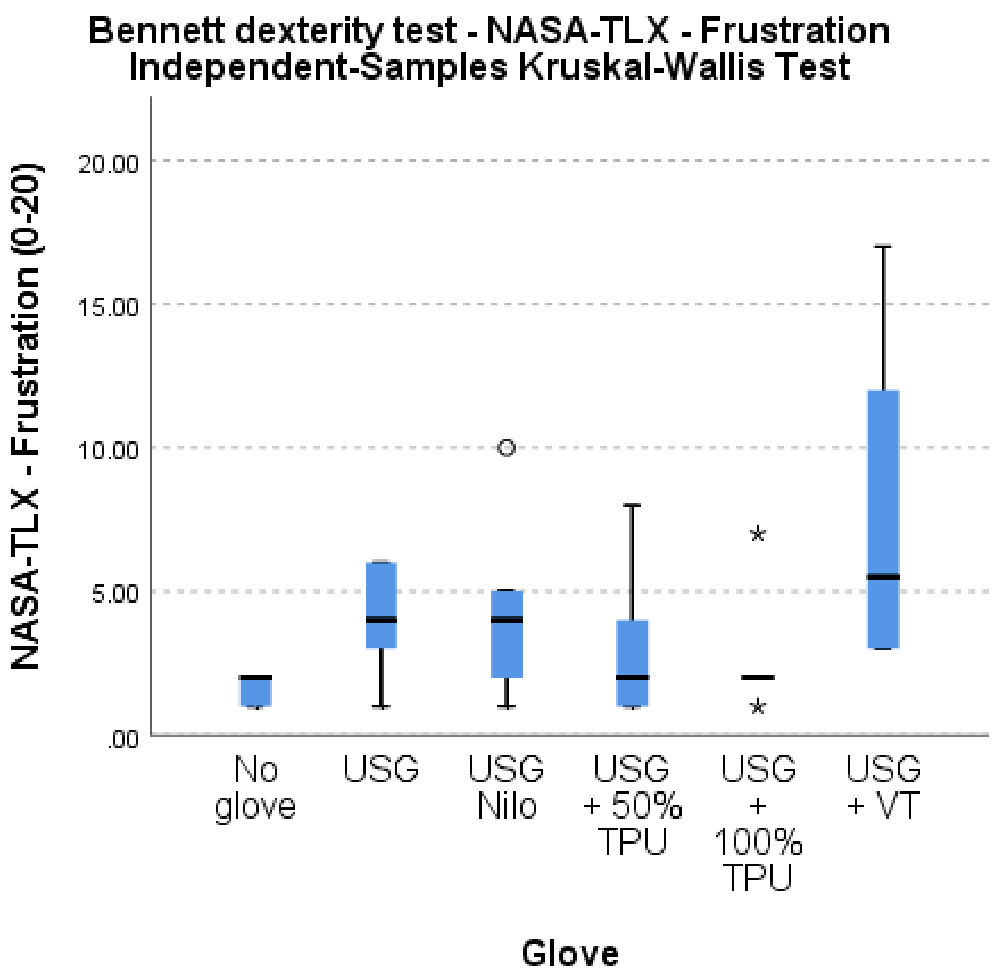
Bennett dexterity test - NASA-TLX - Physical Demand (0-20) field is ordinal but is treated as continuous in the test.

D.3.6 Bennett dexterity test - NASA-TLX - Frustration

Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	11.571[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.041

a. The test statistic is adjusted for ties.



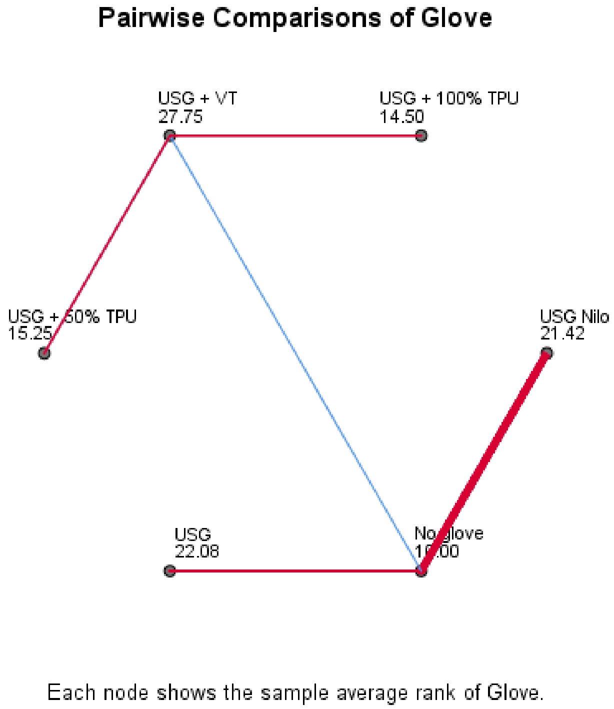
Pairwise Comparisons of Glove

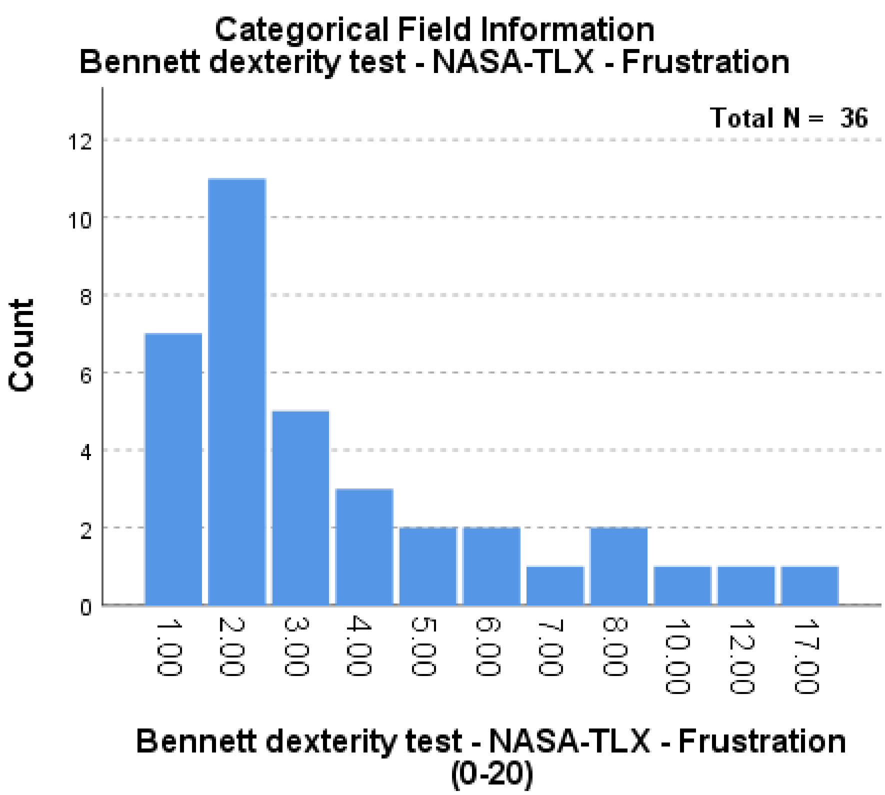
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
No glove-USG + 100% TPU	-4.500	5.963	-.755	.450	1.000
No glove-USG + 50% TPU	-5.250	5.963	-.880	.379	1.000
No glove-USG Nilo	-11.417	5.963	-1.915	.056	.833
No glove-USG	-12.083	5.963	-2.026	.043	.641
No glove-USG + VT	-17.750	5.963	-2.977	.003	.044
USG + 100% TPU-USG + 50% TPU	.750	5.963	.126	.900	1.000
USG + 100% TPU-USG Nilo	6.917	5.963	1.160	.246	1.000
USG + 100% TPU-USG	7.583	5.963	1.272	.203	1.000
USG + 100% TPU-USG + VT	-13.250	5.963	-2.222	.026	.394
USG + 50% TPU-USG Nilo	6.167	5.963	1.034	.301	1.000
USG + 50% TPU-USG	6.833	5.963	1.146	.252	1.000
USG + 50% TPU-USG + VT	-12.500	5.963	-2.096	.036	.541
USG Nilo-USG	.667	5.963	.112	.911	1.000
USG Nilo-USG + VT	-6.333	5.963	-1.062	.288	1.000
USG-USG + VT	-5.667	5.963	-.950	.342	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.





Bennett dexterity test - NASA-TLX - Frustration (0-20) field is ordinal but is treated as continuous in the test.

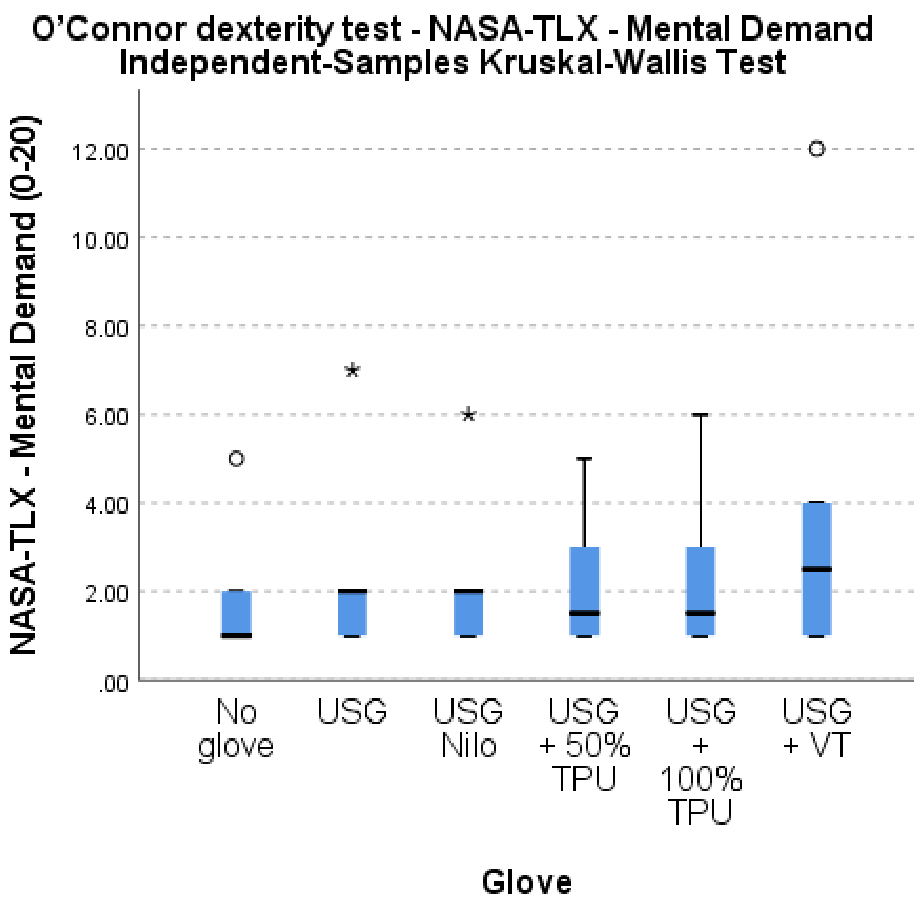
D.3.7 O'Connor dexterity test - NASA-TLX - Mental Demand

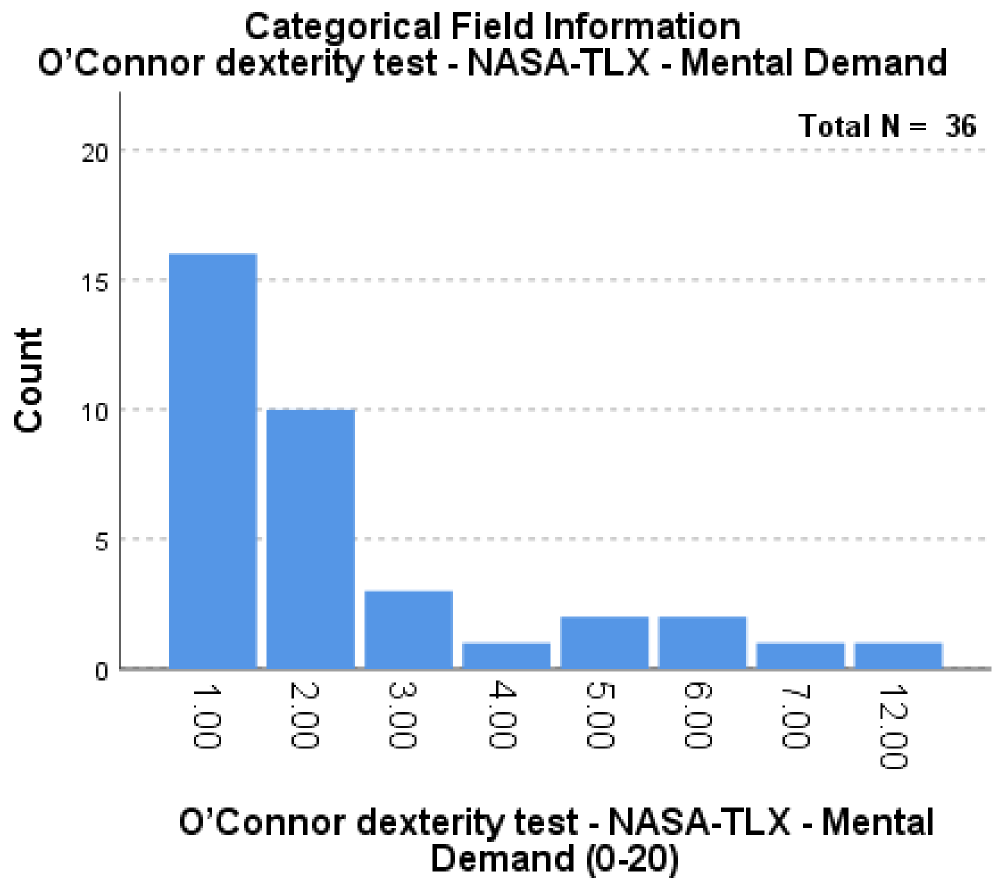
Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	1.873[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.866

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





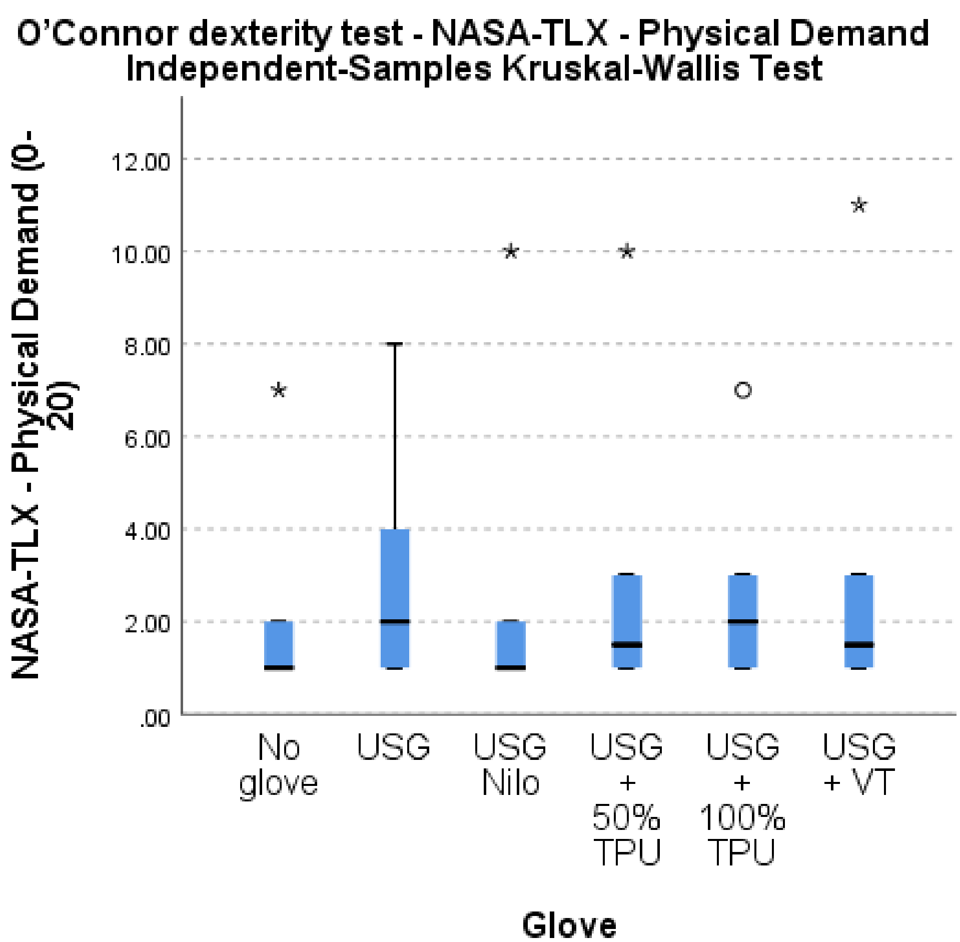
O'Connor dexterity test - NASA-TLX - Mental Demand (0-20) field is ordinal but is treated as continuous in the test.

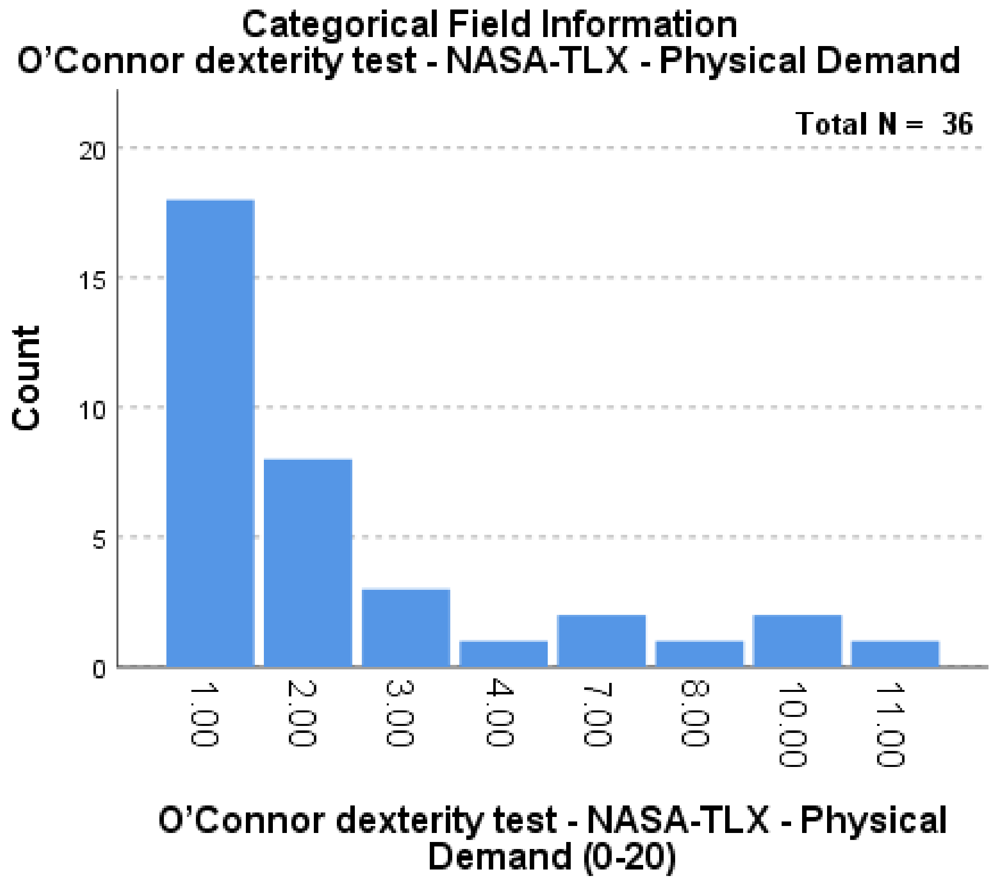
D.3.8 O'Connor dexterity test - NASA-TLX - Physical Demand

Independent-Samples Kruskal-Wallis Test Summary	
Total N	36
Test Statistic	1.828[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.872

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





O'Connor dexterity test - NASA-TLX - Physical Demand (0-20) field is ordinal but is treated as continuous in the test.

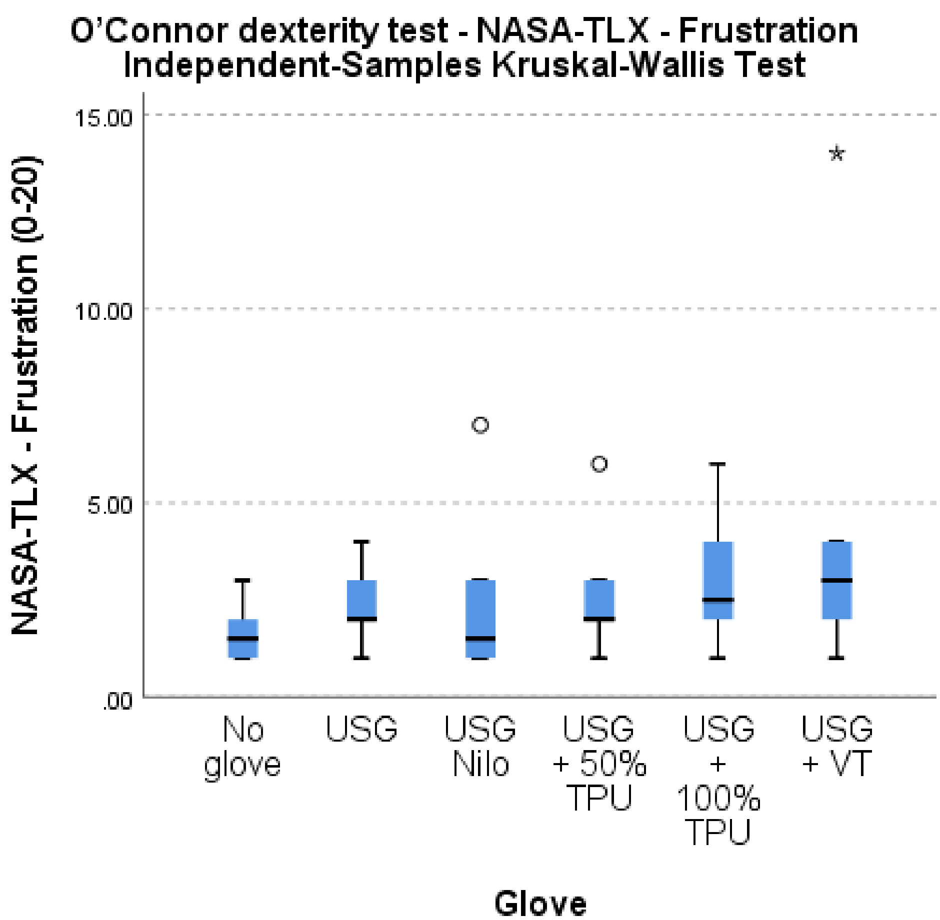
D.3.9 O'Connor dexterity test - NASA-TLX - Frustration

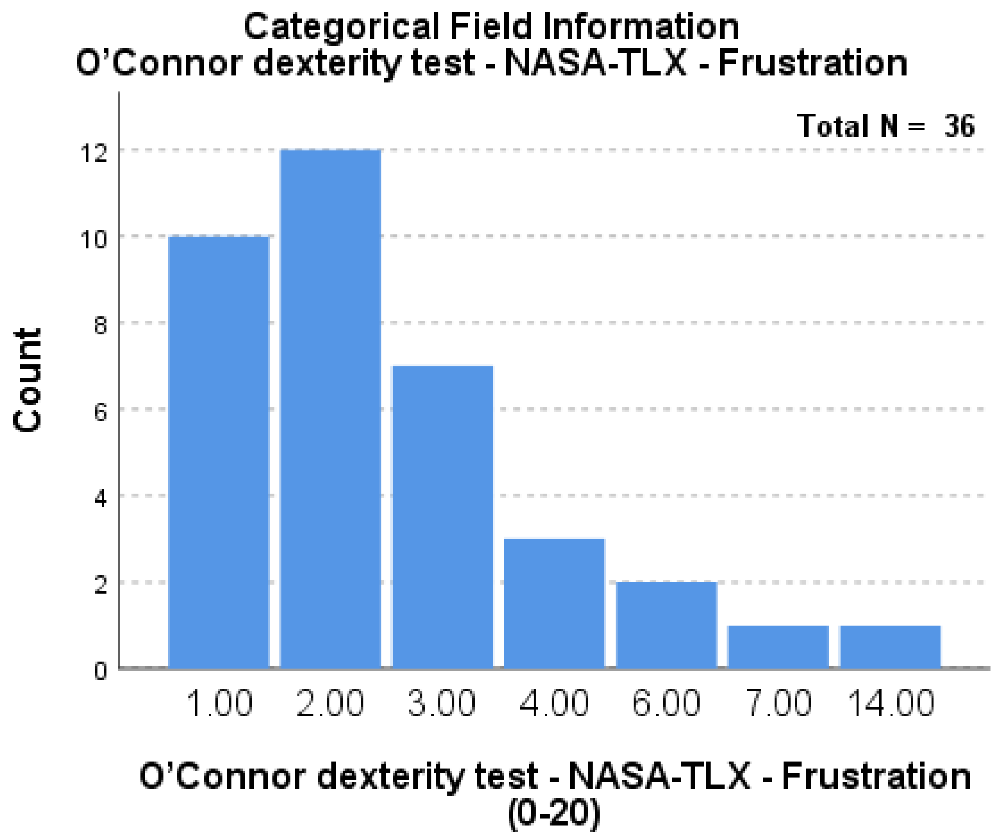
Independent-Samples Kruskal-Wallis Test Summary

Total N	36
Test Statistic	4.479[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.483

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





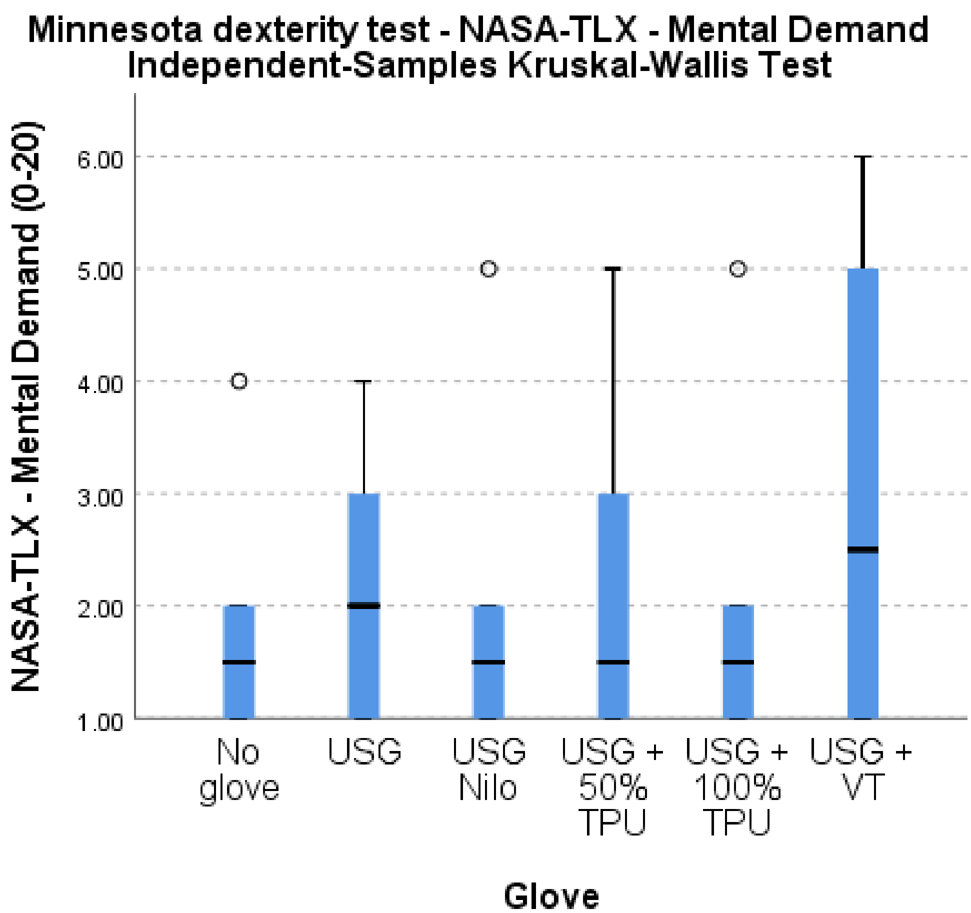
O'Connor dexterity test - NASA-TLX - Frustration (0-20) field is ordinal but is treated as continuous in the test.

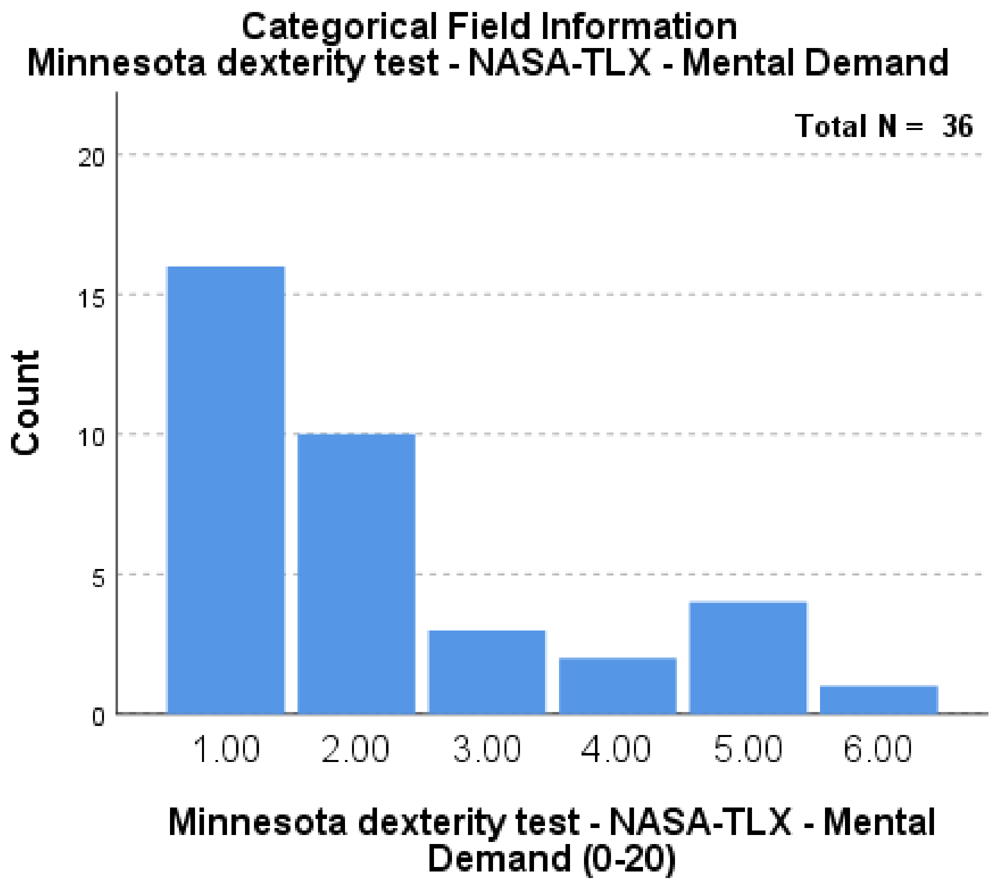
D.3.10 Minnesota dexterity test - NASA-TLX - Mental Demand

Independent-Samples Kruskal-Wallis Test Summary	
Total N	36
Test Statistic	1.677[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.892

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





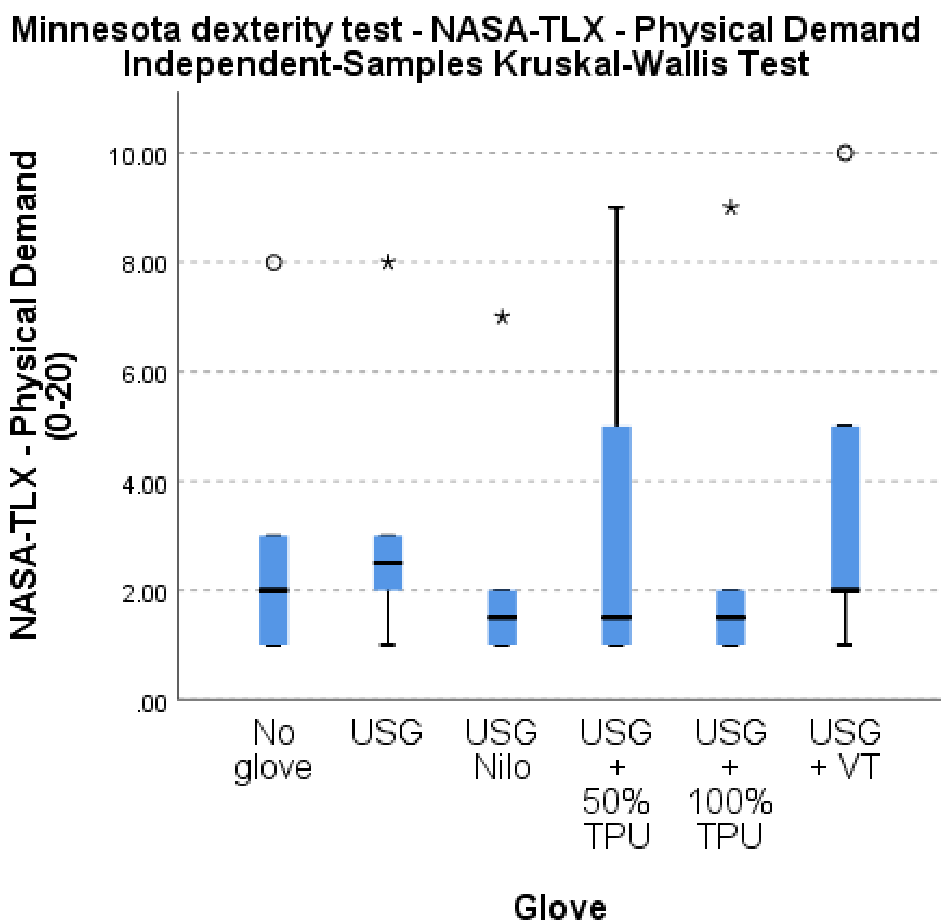
Minnesota dexterity test - NASA-TLX - Mental Demand (0-20) field is ordinal but is treated as continuous in the test.

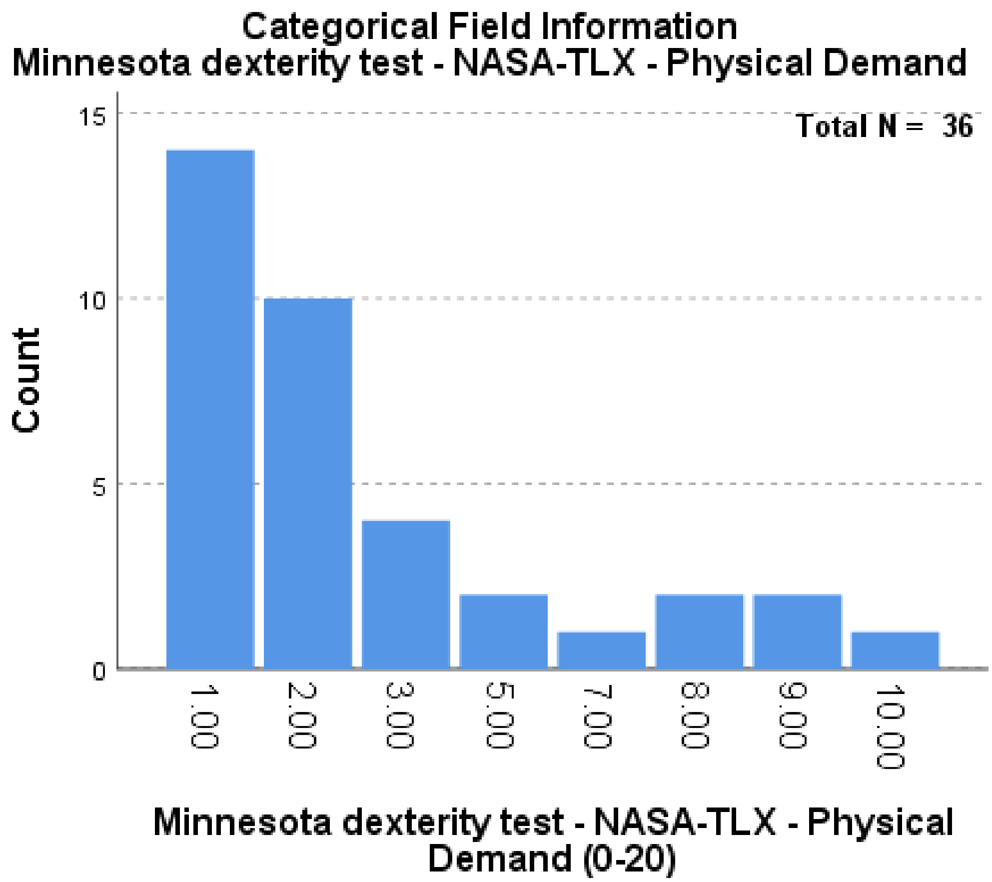
D.3.11 Minnesota dexterity test - NASA-TLX - Physical Demand

Independent-Samples Kruskal-Wallis Test Summary	
Total N	36
Test Statistic	2.391[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.793

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





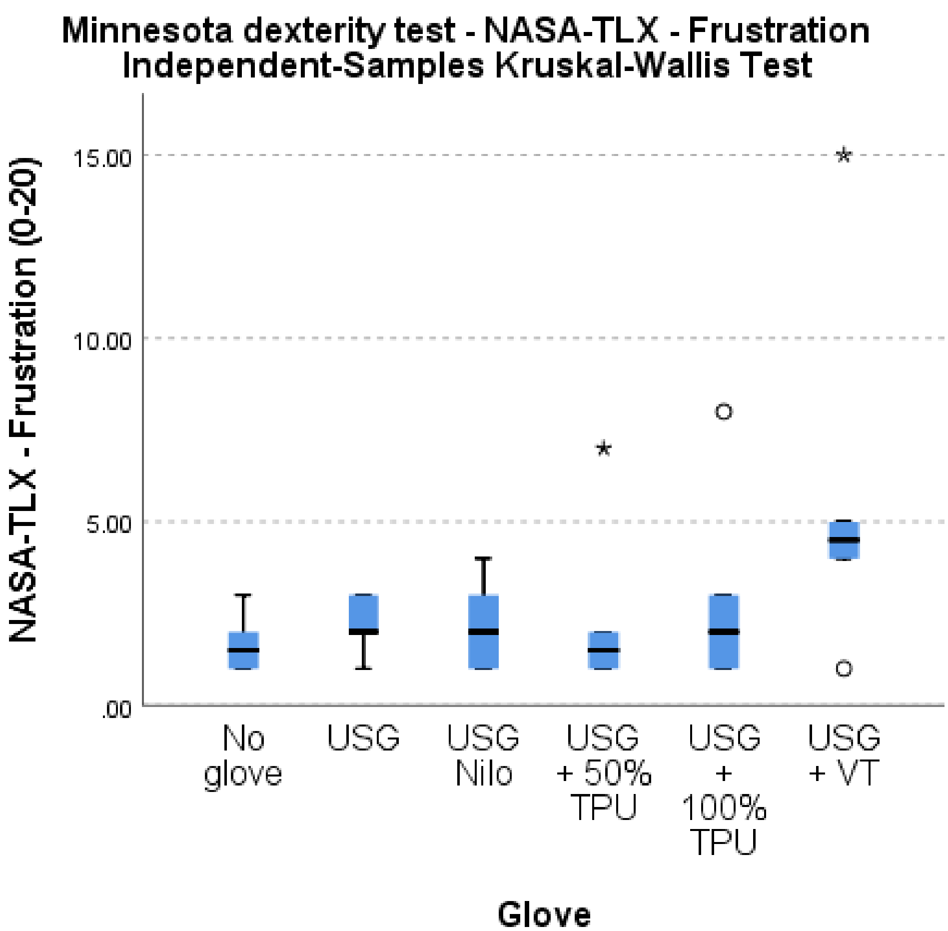
Minnesota dexterity test - NASA-TLX - Physical Demand (0-20) field is ordinal but is treated as continuous in the test.

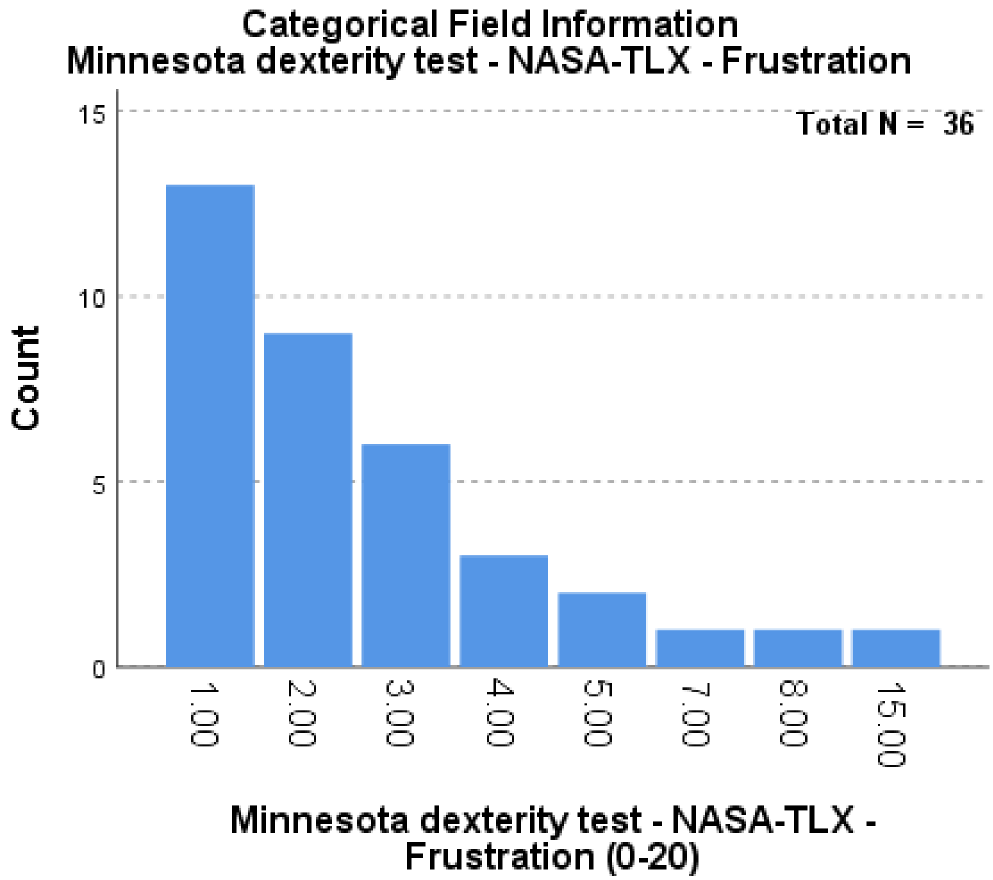
D.3.12 Minnesota dexterity test - NASA-TLX - Frustration

Independent-Samples Kruskal-Wallis Test Summary	
Total N	36
Test Statistic	7.259[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.202

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





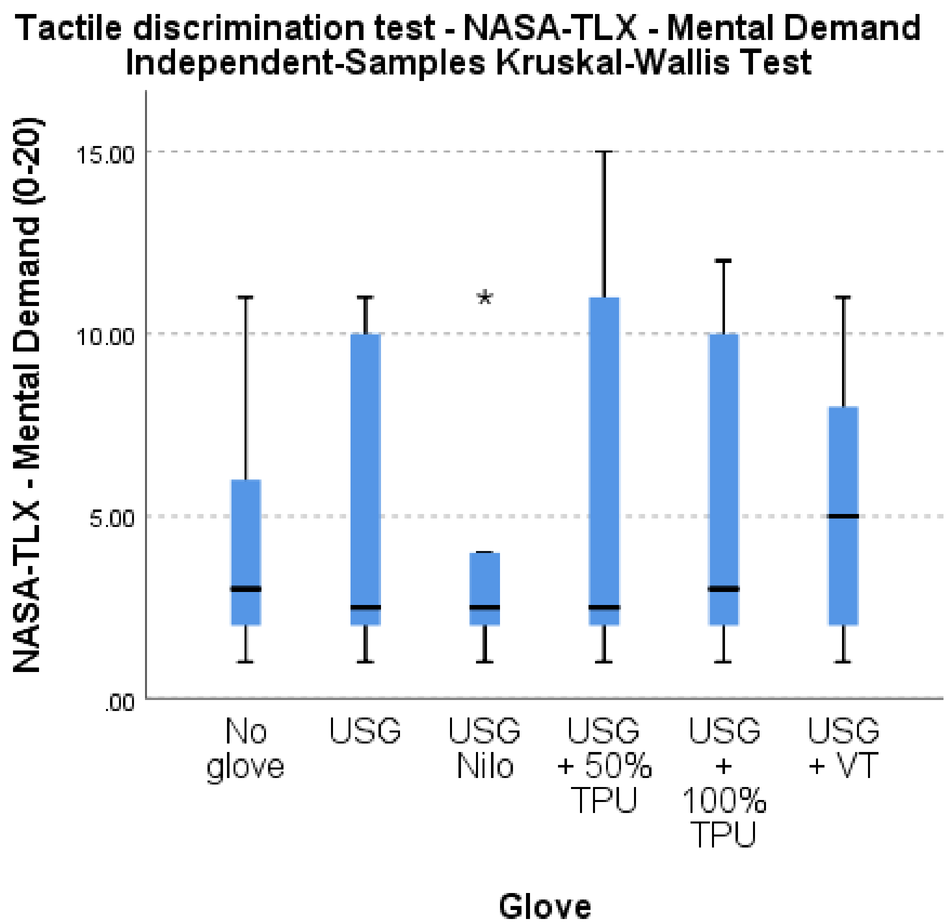
Minnesota dexterity test - NASA-TLX - Frustration (0-20) field is ordinal but is treated as continuous in the test.

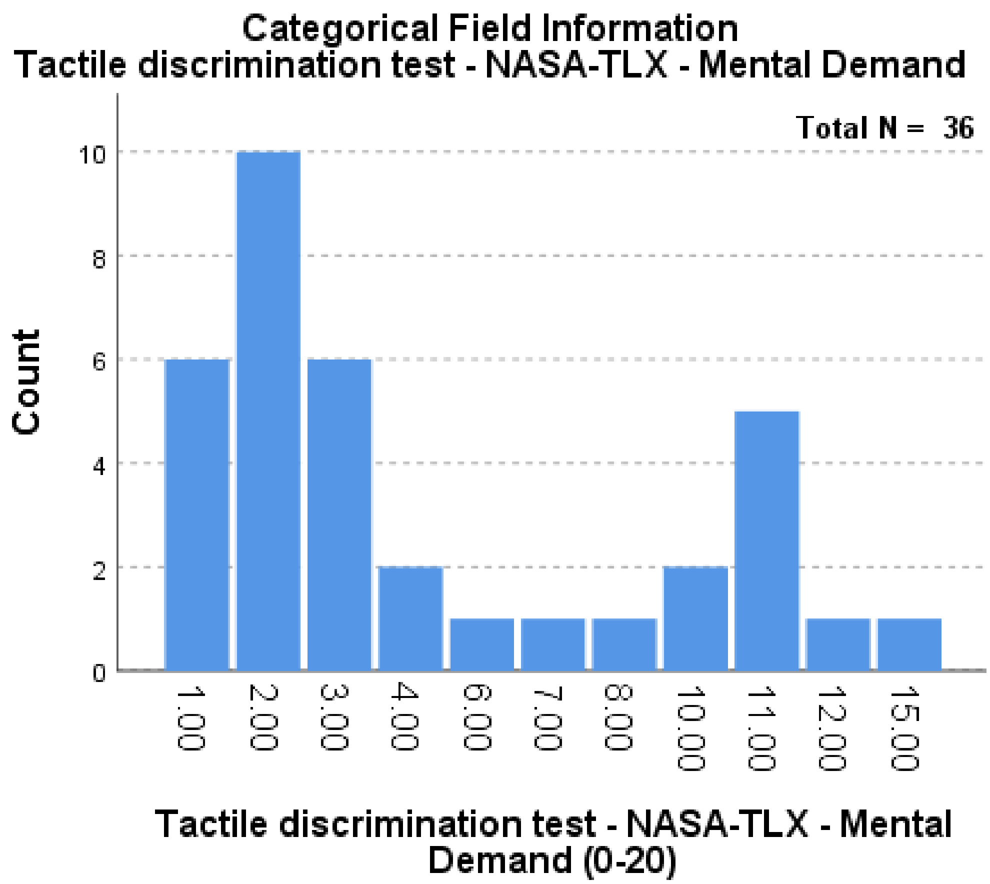
D.3.13 Tactile discrimination test - NASA-TLX - Mental Demand

Independent-Samples Kruskal-Wallis Test Summary	
Total N	36
Test Statistic	.307[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.998

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





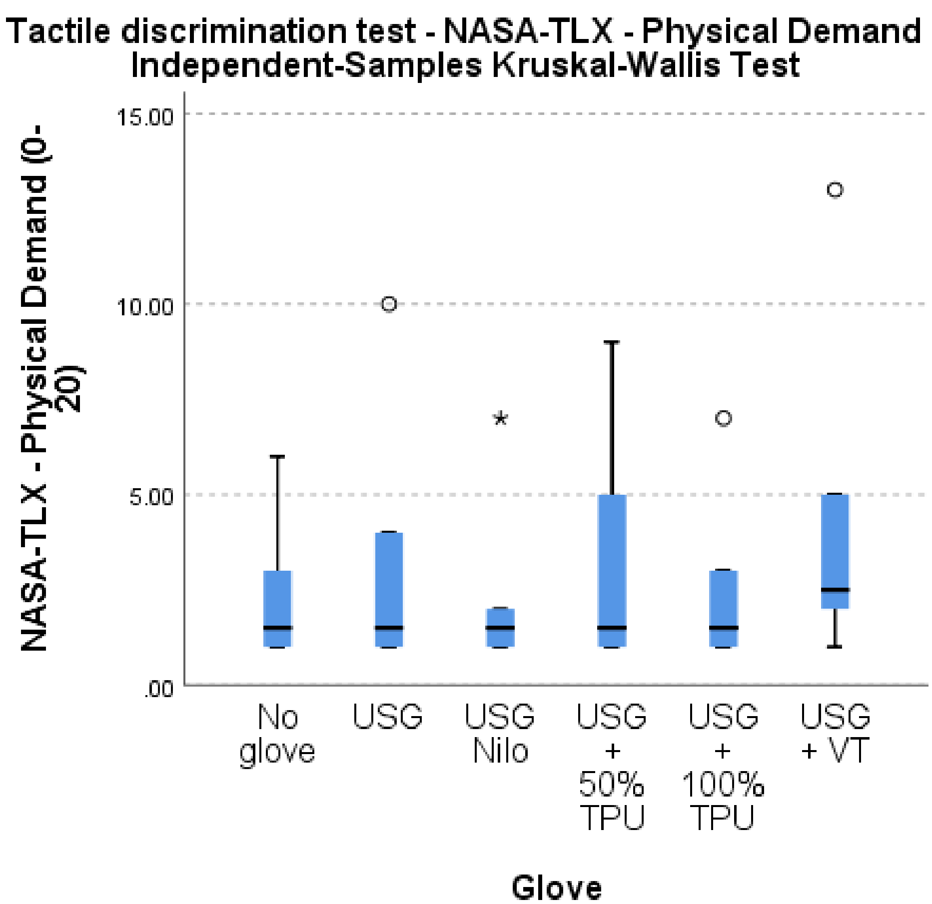
Tactile discrimination test - NASA-TLX - Mental Demand (0-20) field is ordinal but is treated as continuous in the test.

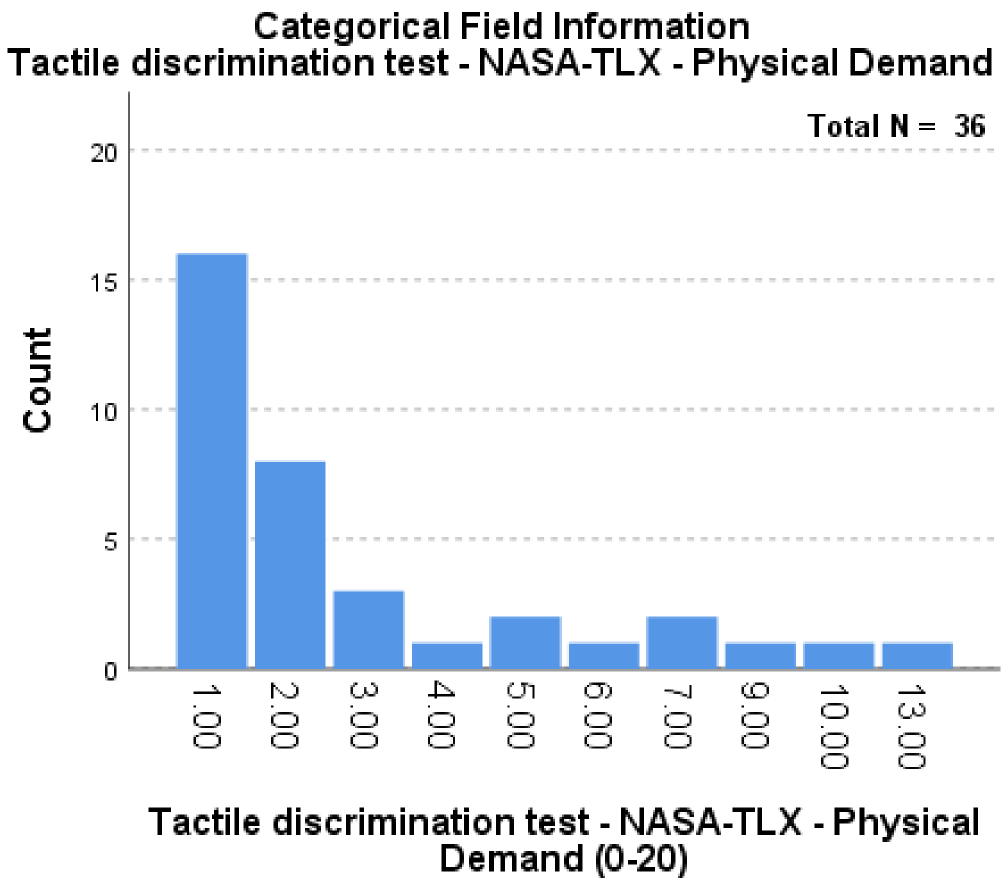
D.3.14 Tactile discrimination test - NASA-TLX - Physical Demand

Independent-Samples Kruskal-Wallis Test Summary	
Total N	36
Test Statistic	1.928[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.859

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





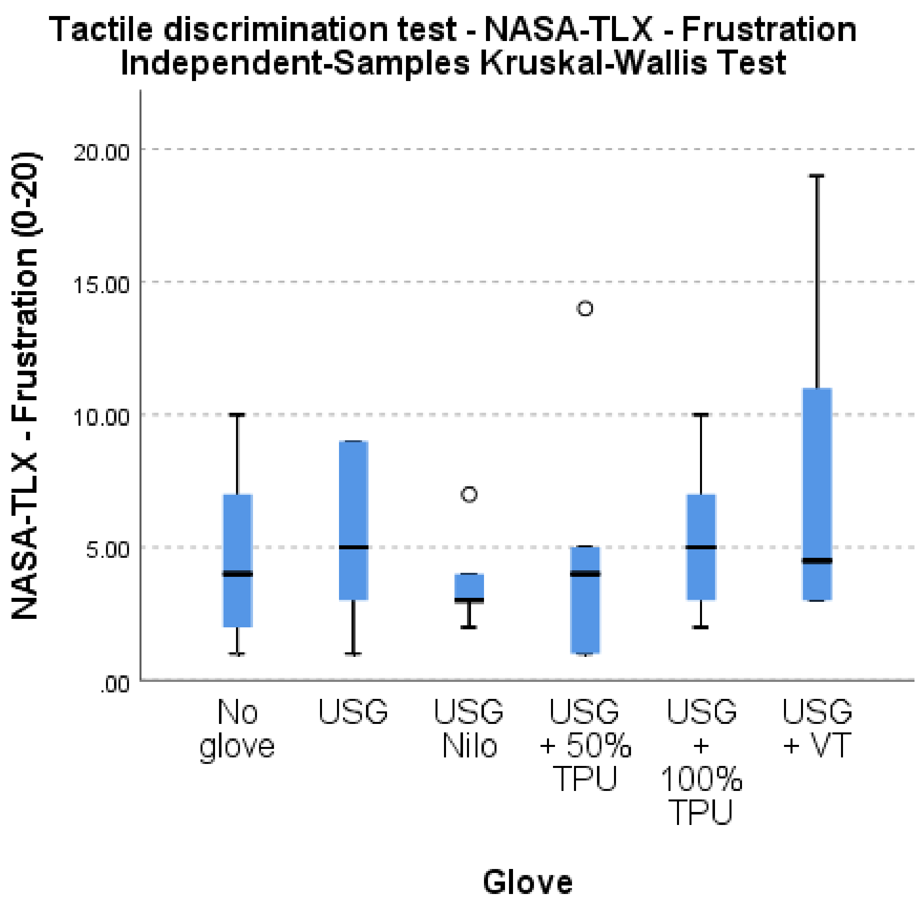
Tactile discrimination test - NASA-TLX - Physical Demand (0-20) field is ordinal but is treated as continuous in the test.

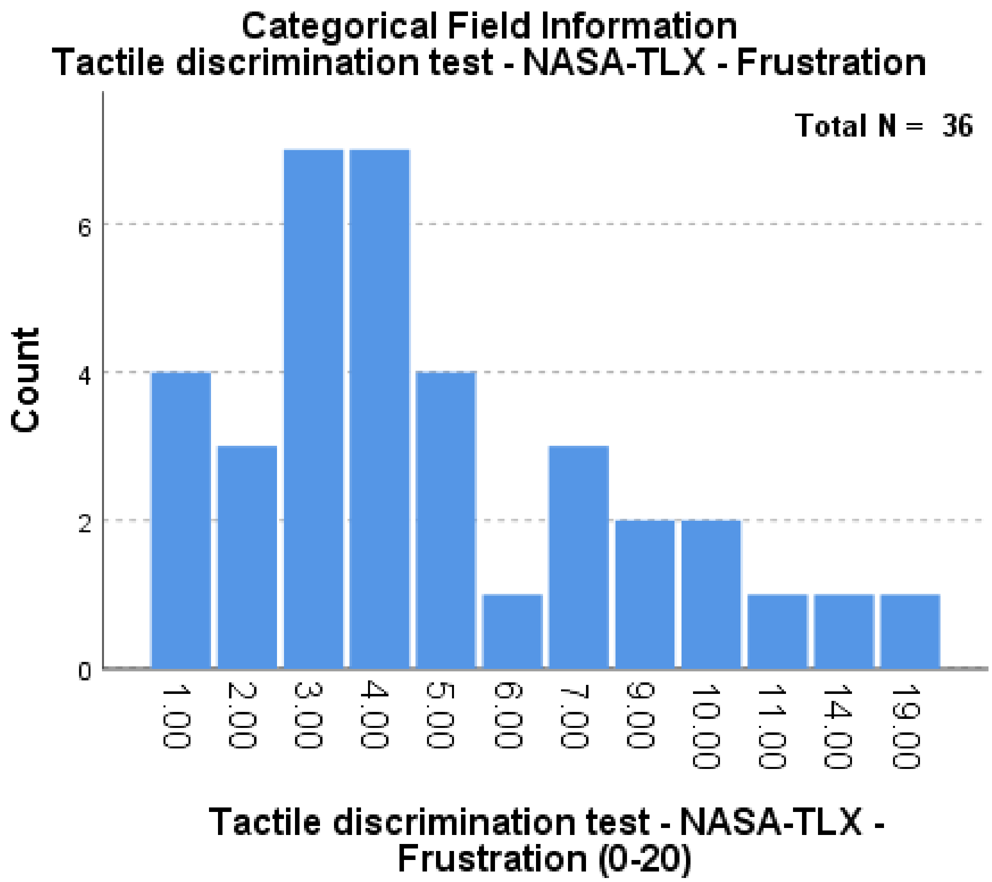
D.3.15 Tactile discrimination test - NASA-TLX - Frustration

Independent-Samples Kruskal-Wallis Test Summary	
Total N	36
Test Statistic	2.430[a,b]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.787

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





Tactile discrimination test - NASA-TLX - Frustration (0-20) field is ordinal but is treated as continuous in the test.

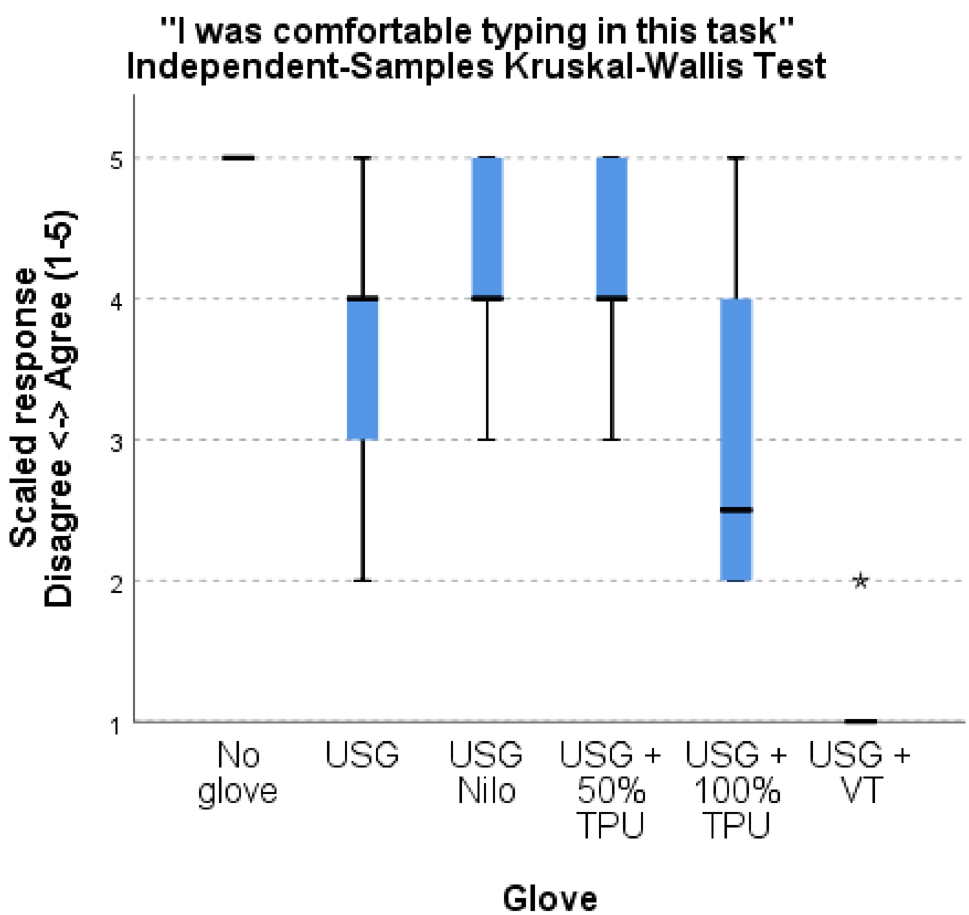
D.4 Scaled response data

D.4.1 "I was comfortable typing in this task"

Independent-Samples Kruskal-Wallis Test Summary

Total N	35
Test Statistic	21.443[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.001

a. The test statistic is adjusted for ties.



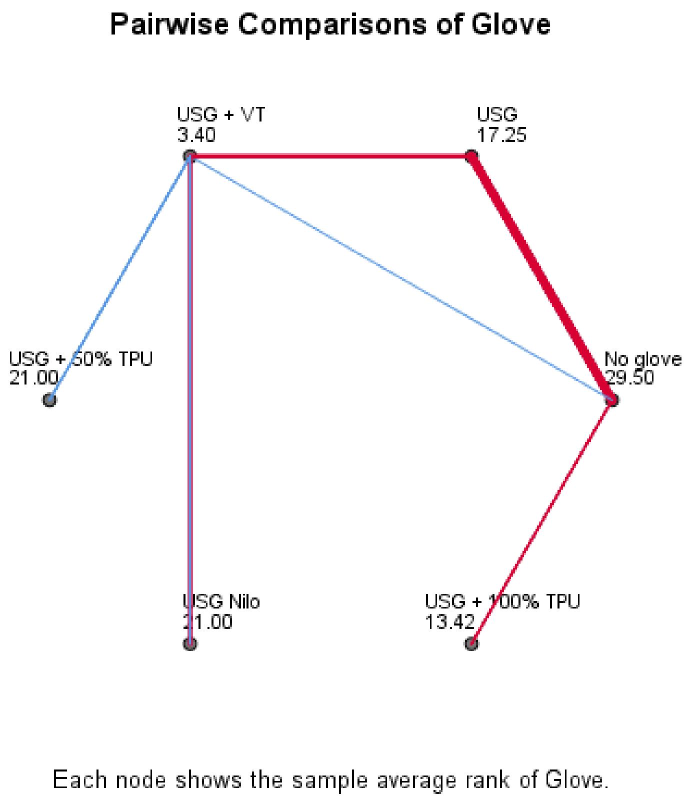
Pairwise Comparisons of Glove

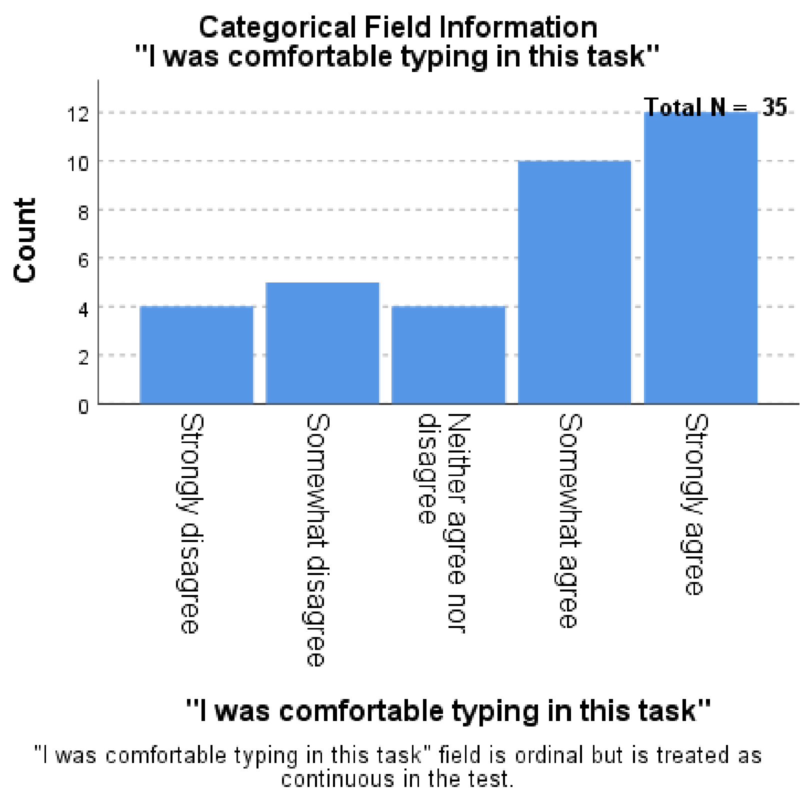
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
USG + VT-USG + 100% TPU	10.017	5.988	1.673	.094	1.000
USG + VT-USG	13.850	5.988	2.313	.021	.311
USG + VT-USG Nilo	17.600	5.988	2.939	.003	.049
USG + VT-USG + 50% TPU	17.600	5.988	2.939	.003	.049
USG + VT-No glove	26.100	5.988	4.359	.000	.000
USG + 100% TPU-USG	3.833	5.709	.671	.502	1.000
USG + 100% TPU-USG Nilo	7.583	5.709	1.328	.184	1.000
USG + 100% TPU-USG + 50% TPU	7.583	5.709	1.328	.184	1.000
USG + 100% TPU-No glove	16.083	5.709	2.817	.005	.073
USG-USG Nilo	-3.750	5.709	-.657	.511	1.000
USG-USG + 50% TPU	-3.750	5.709	-.657	.511	1.000
USG-No glove	12.250	5.709	2.146	.032	.478
USG Nilo-No glove	8.500	5.709	1.489	.137	1.000
USG + 50% TPU-No glove	8.500	5.709	1.489	.137	1.000
USG Nilo-USG + 50% TPU	.000	5.709	.000	1.000	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.


Pairwise Comparisons of Glove
USG + VT
3.40
USG
17.25
USG + 50% TPU
21.00
No glove
29.50
USG Nilo
21.00
USG + 100% TPU
13.42
Each node shows the sample average rank of Glove.

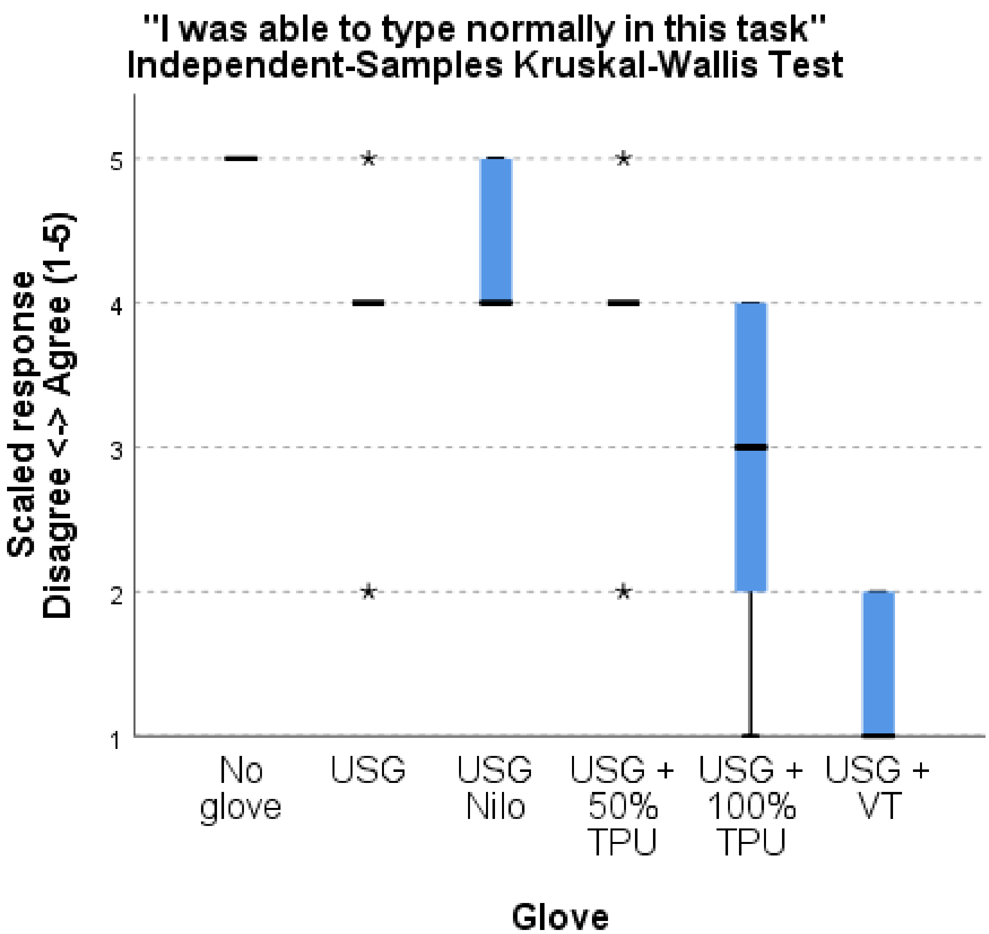

Categorical Field Information
"I was comfortable typing in this task"
Total N = 35
Count
0
2
4
6
8
10
12
Strongly disagree
Somewhat disagree
Neither agree nor disagree
Somewhat agree
Strongly agree
"I was comfortable typing in this task"
"I was comfortable typing in this task" field is ordinal but is treated as continuous in the test.

D.4.2 "I was able to type normally in this task"

Independent-Samples Kruskal-Wallis Test Summary

Total N	35
Test Statistic	23.232[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

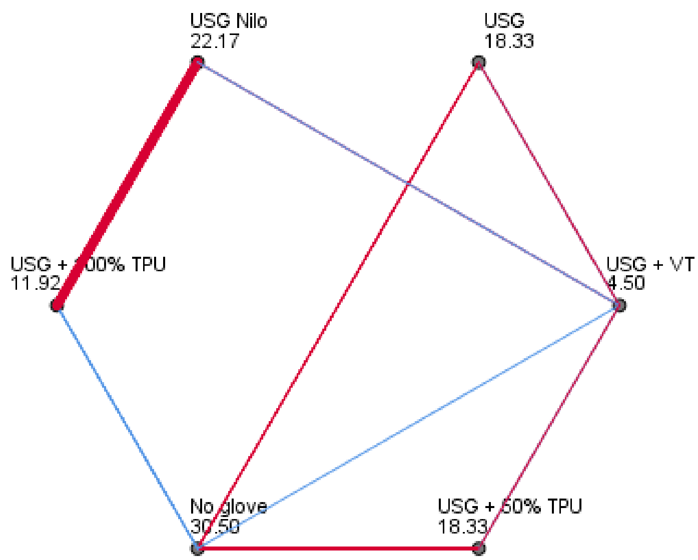
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
USG + VT-USG + 100% TPU	7.417	5.861	1.265	.206	1.000
USG + VT-USG	13.833	5.861	2.360	.018	.274
USG + VT-USG + 50% TPU	13.833	5.861	2.360	.018	.274
USG + VT-USG Nilo	17.667	5.861	3.014	.003	.039
USG + VT-No glove	26.000	5.861	4.436	.000	.000
USG + 100% TPU-USG	6.417	5.588	1.148	.251	1.000
USG + 100% TPU-USG + 50% TPU	6.417	5.588	1.148	.251	1.000
USG + 100% TPU-USG Nilo	10.250	5.588	1.834	.067	.999
USG + 100% TPU-No glove	18.583	5.588	3.326	.001	.013
USG-No glove	12.167	5.588	2.177	.029	.442
USG + 50% TPU-No glove	12.167	5.588	2.177	.029	.442
USG-USG + 50% TPU	.000	5.588	.000	1.000	1.000
USG-USG Nilo	-3.833	5.588	-.686	.493	1.000
USG + 50% TPU-USG Nilo	3.833	5.588	.686	.493	1.000
USG Nilo-No glove	8.333	5.588	1.491	.136	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

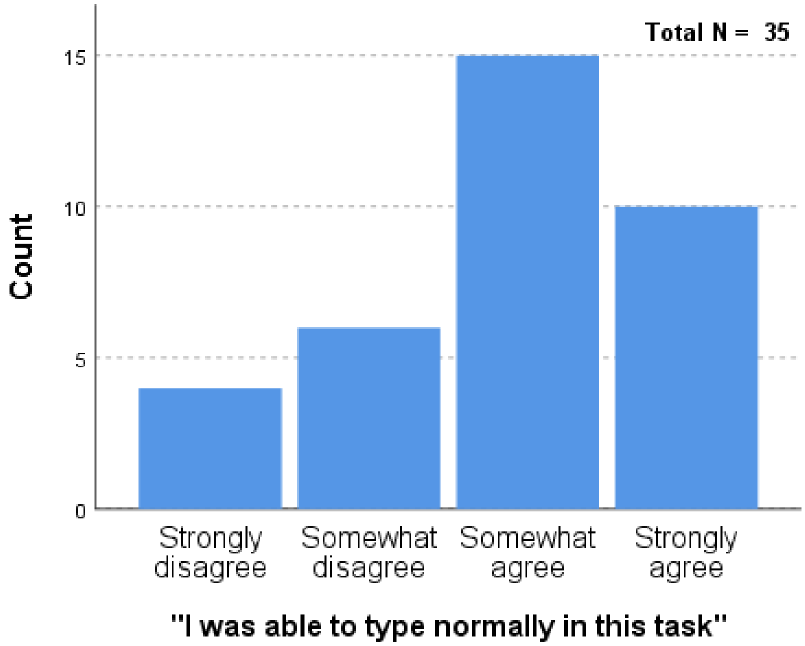
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.

Categorical Field Information
"I was able to type normally in this task"

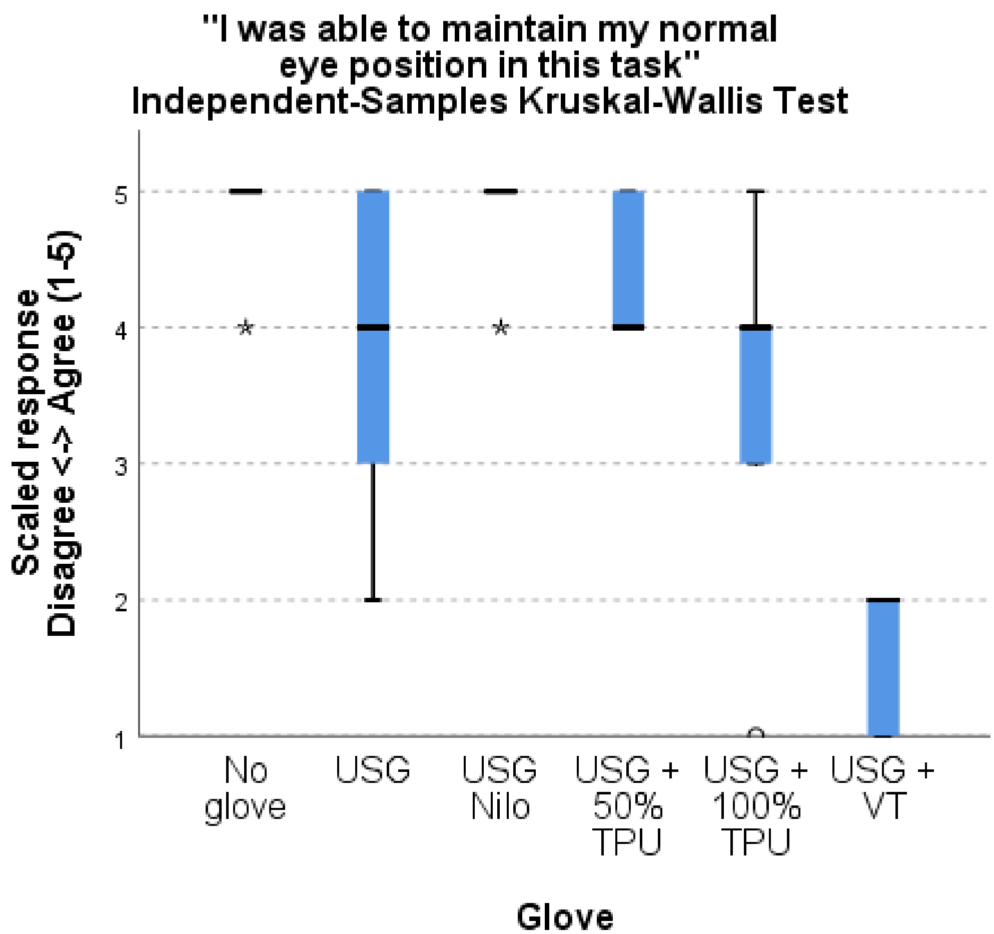


"I was able to type normally in this task" field is ordinal but is treated as continuous in the test.

D.4.3 "I was able to maintain my normal eye position in this task"

Independent-Samples Kruskal-Wallis Test Summary	
Total N	35
Test Statistic	19.545[a]
Degree Of Freedom	5
Asymptotic Sig.(2-sided test)	.002

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

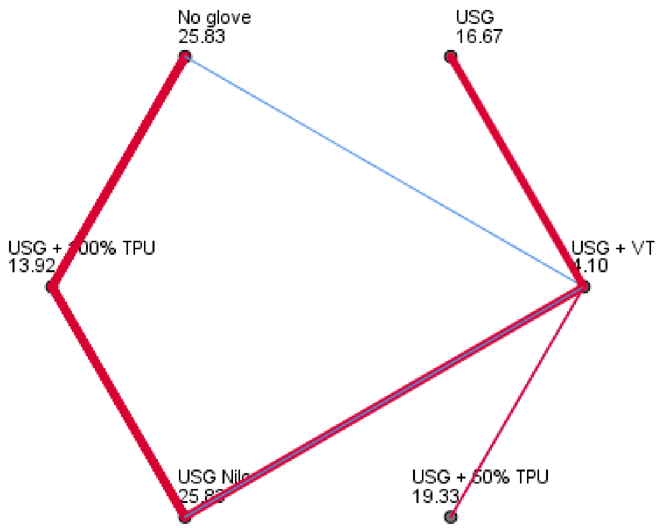
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
USG + VT-USG + 100% TPU	9.817	5.849	1.678	.093	1.000
USG + VT-USG	12.567	5.849	2.148	.032	.475
USG + VT-USG + 50% TPU	15.233	5.849	2.604	.009	.138
USG + VT-No glove	21.733	5.849	3.716	.000	.003
USG + VT-USG Nilo	21.733	5.849	3.716	.000	.003
USG + 100% TPU-USG	2.750	5.577	.493	.622	1.000
USG + 100% TPU-USG + 50% TPU	5.417	5.577	.971	.331	1.000
USG + 100% TPU-No glove	11.917	5.577	2.137	.033	.489
USG + 100% TPU-USG Nilo	11.917	5.577	2.137	.033	.489
USG-USG + 50% TPU	-2.667	5.577	-.478	.633	1.000
USG-No glove	9.167	5.577	1.644	.100	1.000
USG-USG Nilo	-9.167	5.577	-1.644	.100	1.000
USG + 50% TPU-No glove	6.500	5.577	1.166	.244	1.000
USG + 50% TPU-USG Nilo	6.500	5.577	1.166	.244	1.000
No glove-USG Nilo	.000	5.577	.000	1.000	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

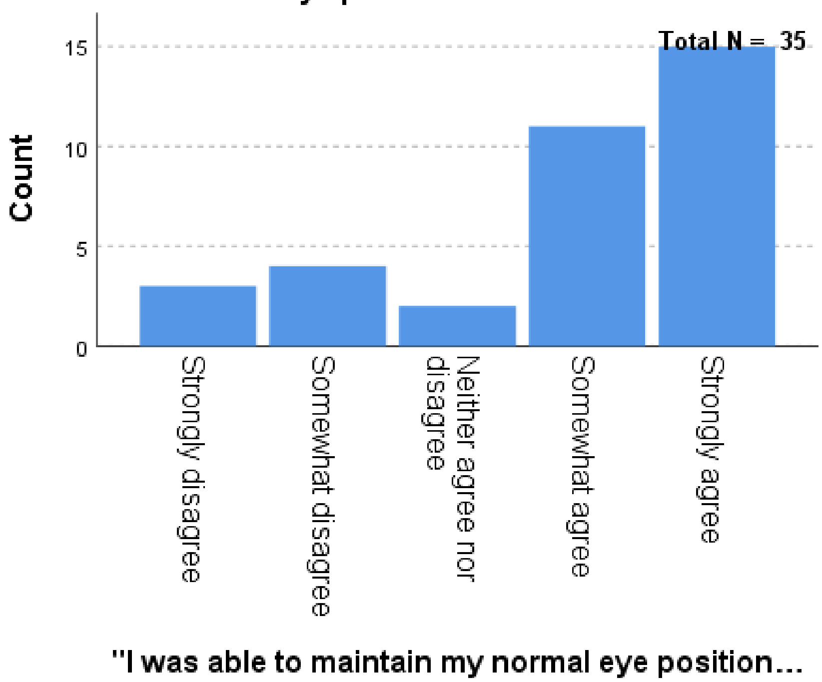
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.

Categorical Field Information "I was able to maintain my normal eye position in this task"



"I was able to maintain my normal eye position in this task" field is ordinal but is treated as continuous in the test.

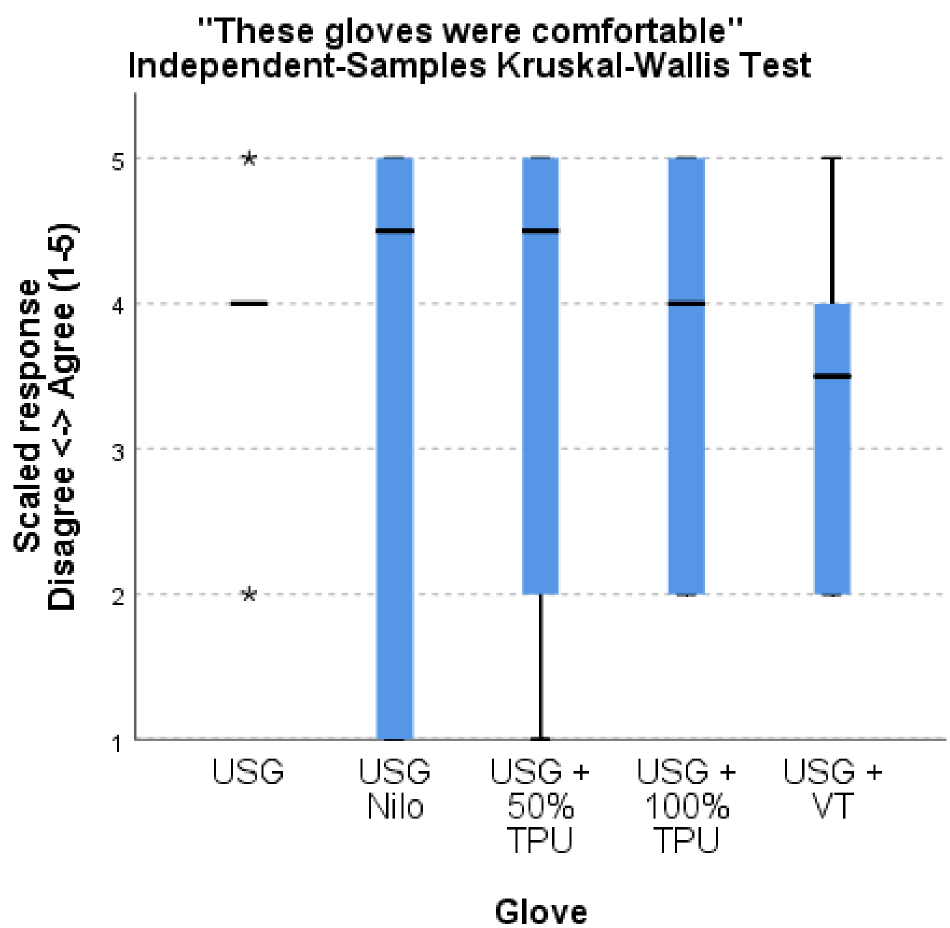
D.4.4 "These gloves were comfortable"

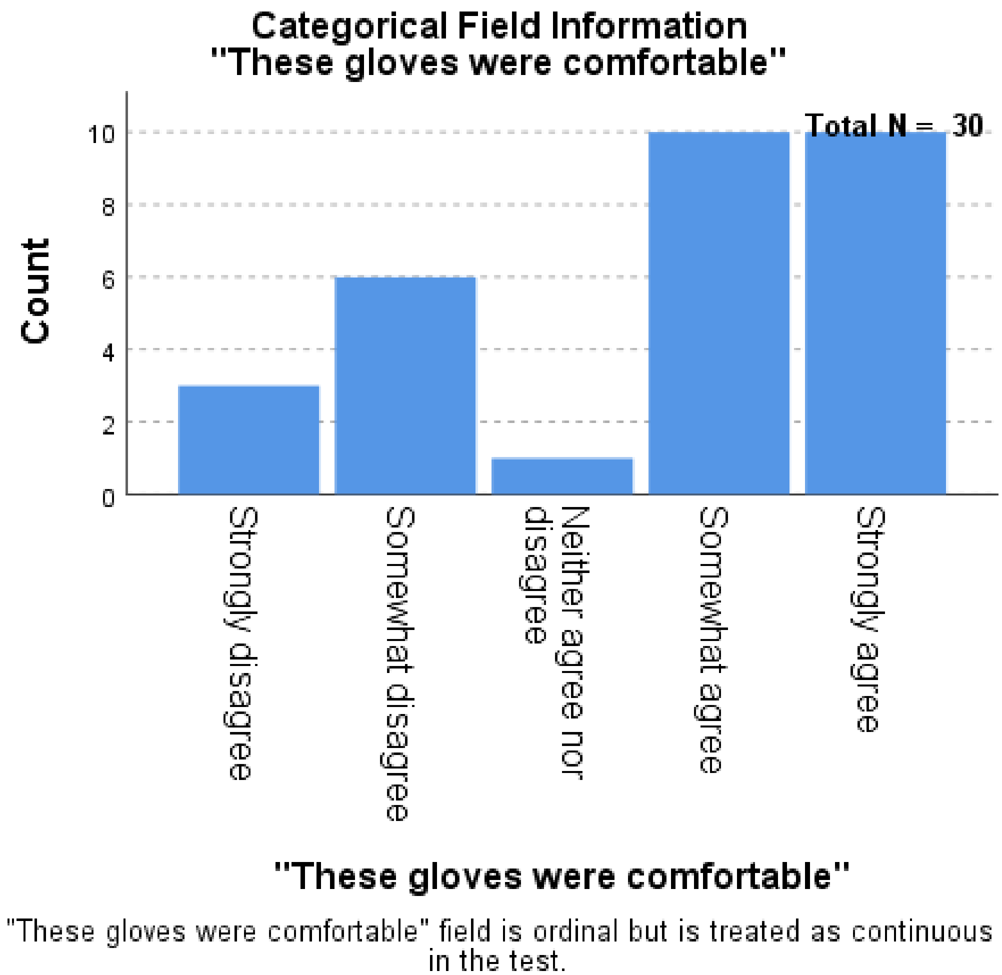
Independent-Samples Kruskal-Wallis Test Summary

Total N	30
Test Statistic	.592[a,b]
Degree Of Freedom	4
Asymptotic Sig.(2-sided test)	.964

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.



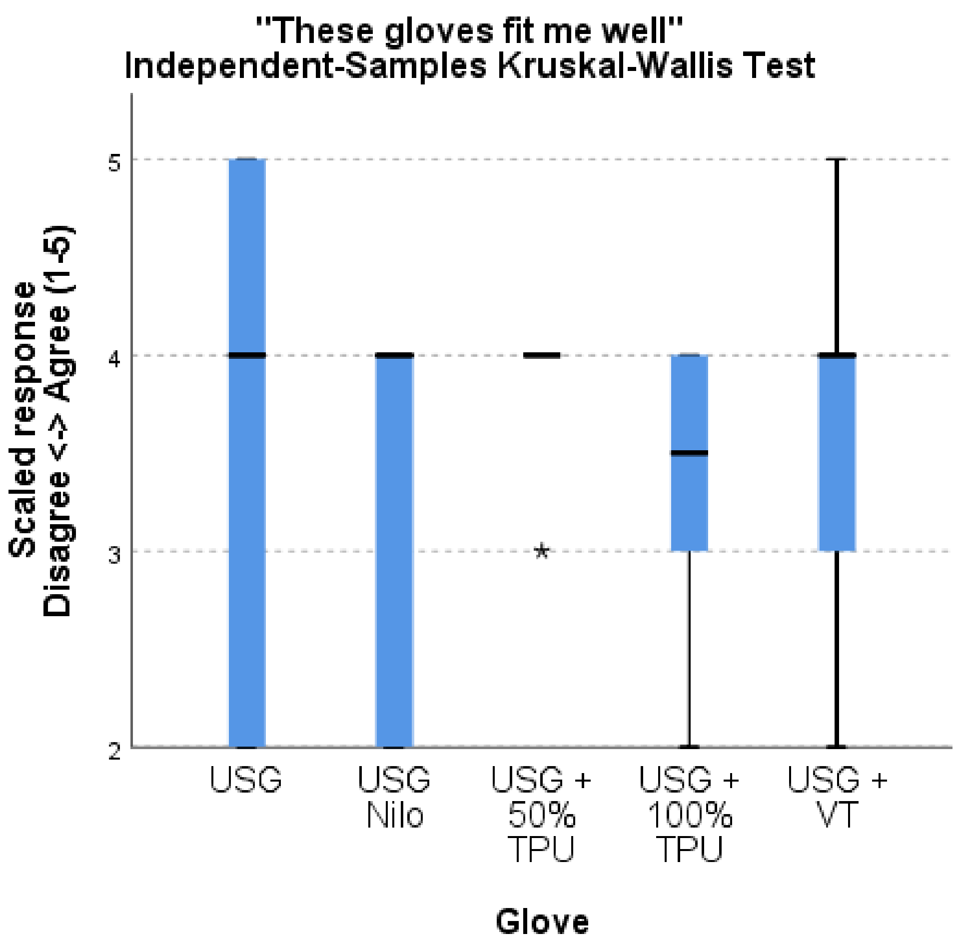

Categorical Field Information
"These gloves were comfortable"
Total N = 30
Count
0
2
4
6
8
10
Strongly disagree
Somewhat disagree
Neither agree nor disagree
Somewhat agree
Strongly agree
"These gloves were comfortable"
"These gloves were comfortable" field is ordinal but is treated as continuous in the test.

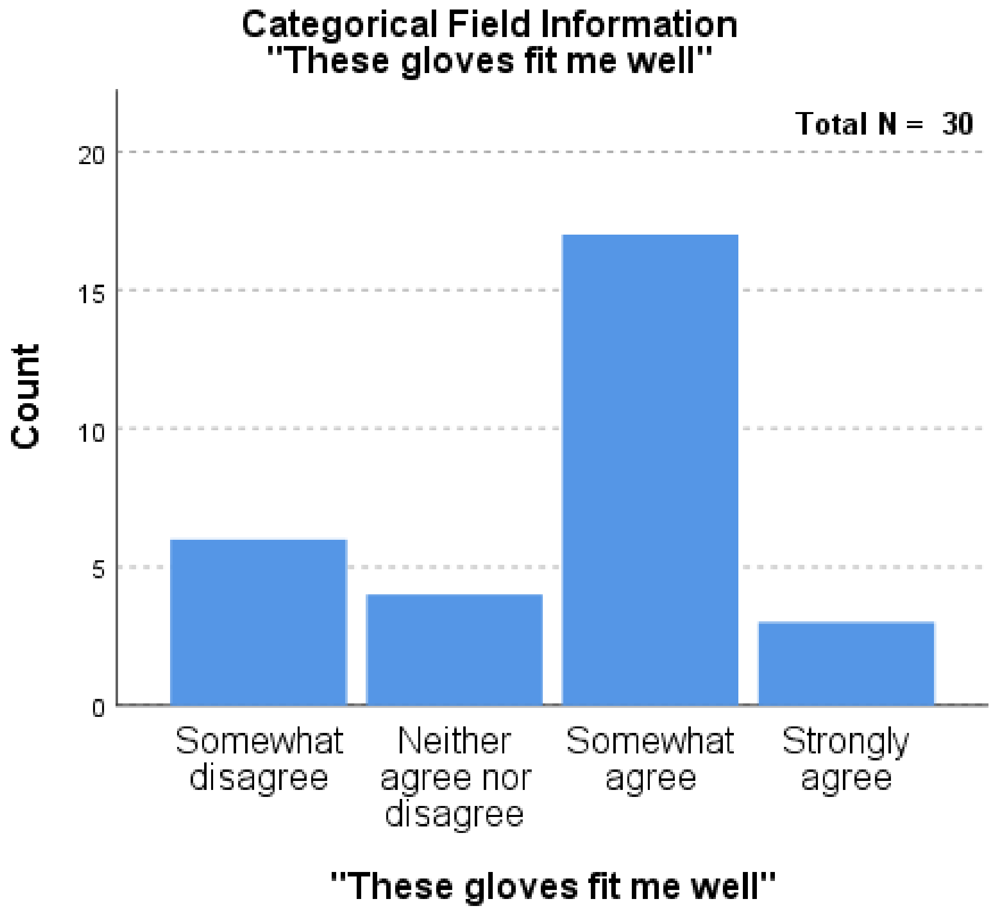
D.4.5 "These gloves fit me well"

Independent-Samples Kruskal-Wallis Test Summary	
Total N	30
Test Statistic	1.533[a,b]
Degree Of Freedom	4
Asymptotic Sig.(2-sided test)	.821

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.





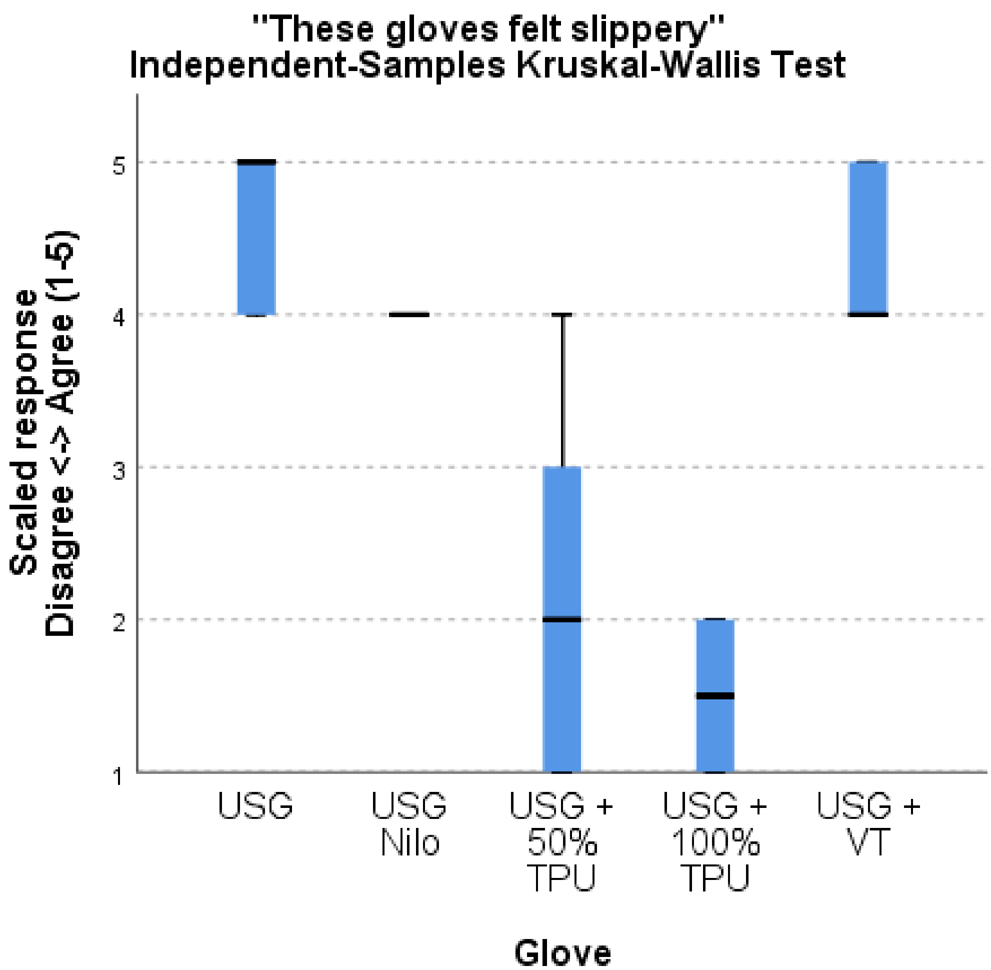
"These gloves fit me well" field is ordinal but is treated as continuous in th...

D.4.6 "These gloves felt slippery"

Independent-Samples Kruskal-Wallis Test Summary

Total N	30
Test Statistic	22.770[a]
Degree Of Freedom	4
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



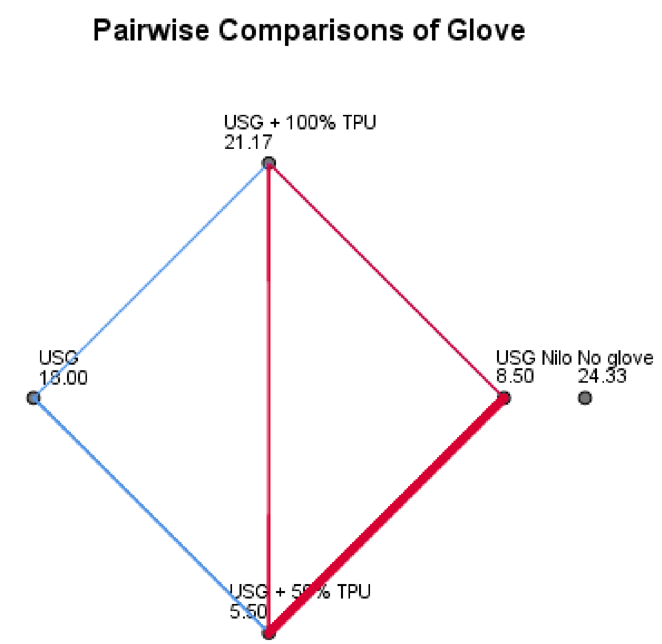
Pairwise Comparisons of Glove

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
USG + 100% TPU-USG + 50% TPU	3.000	4.828	.621	.534	1.000
USG + 100% TPU-USG Nilo	12.500	4.828	2.589	.010	.144
USG + 100% TPU-USG + VT	-15.667	4.828	-3.245	.001	.018
USG + 100% TPU-USG	18.833	4.828	3.901	.000	.001
USG + 50% TPU-USG Nilo	9.500	4.828	1.968	.049	.737
USG + 50% TPU-USG + VT	-12.667	4.828	-2.624	.009	.131
USG + 50% TPU-USG	15.833	4.828	3.279	.001	.016
USG Nilo-USG + VT	-3.167	4.828	-.656	.512	1.000
USG Nilo-USG	6.333	4.828	1.312	.190	1.000
USG + VT-USG	3.167	4.828	.656	.512	1.000

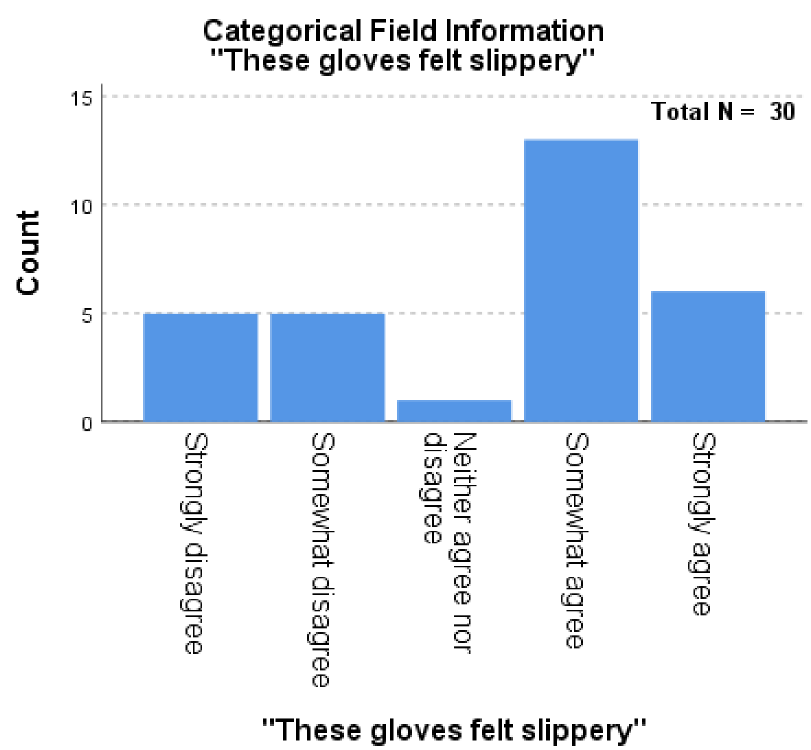
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



Each node shows the sample average rank of Glove.



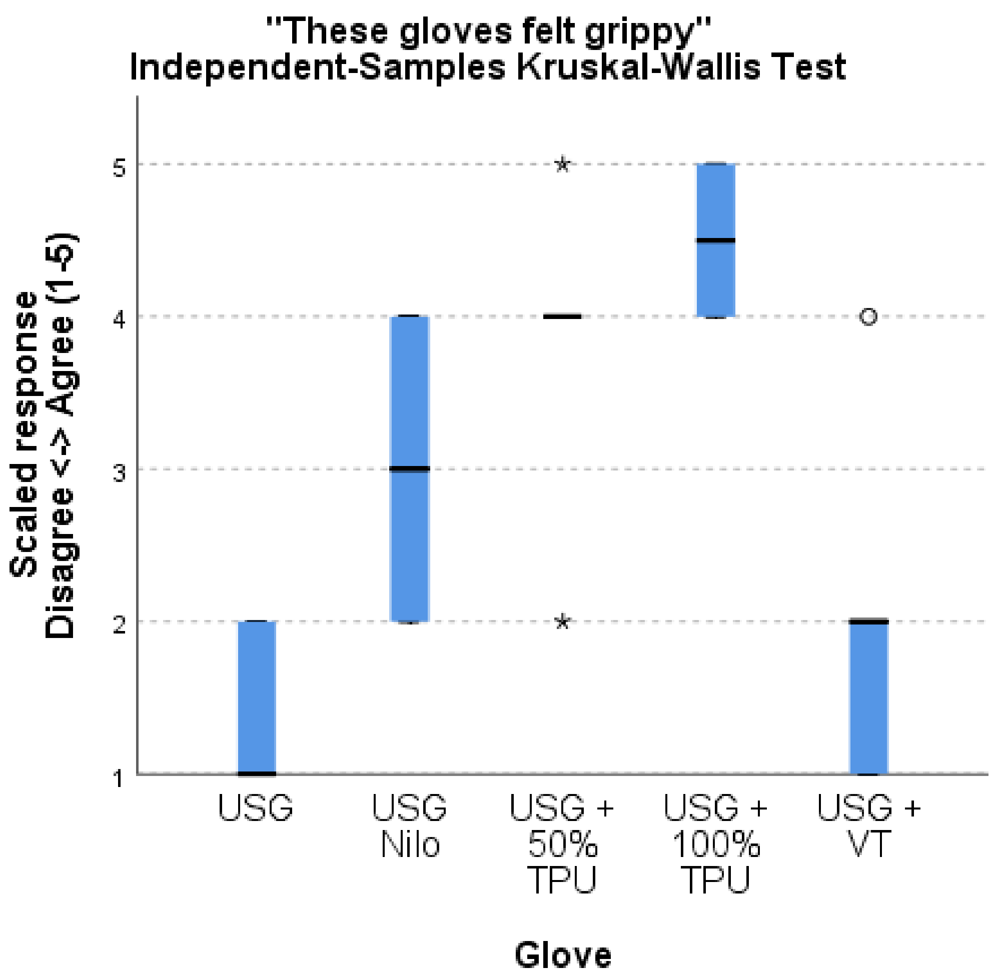
"These gloves felt slippery" field is ordinal but is treated as continuous in t…

D.4.7 "These gloves felt grippy"

Independent-Samples Kruskal-Wallis Test Summary

Total N	30
Test Statistic	20.093[a]
Degree Of Freedom	4
Asymptotic Sig.(2-sided test)	.000

a. The test statistic is adjusted for ties.



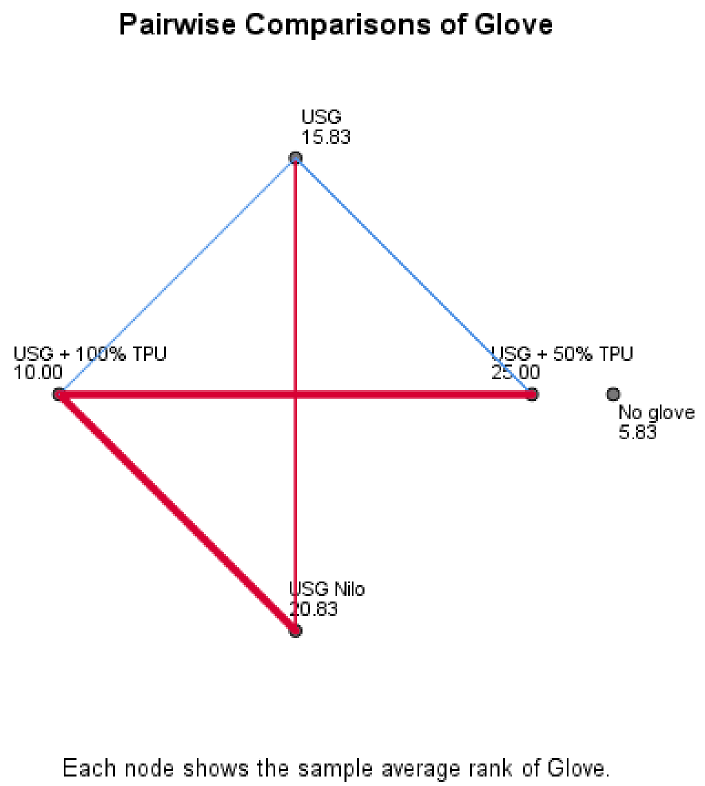
Pairwise Comparisons of Glove

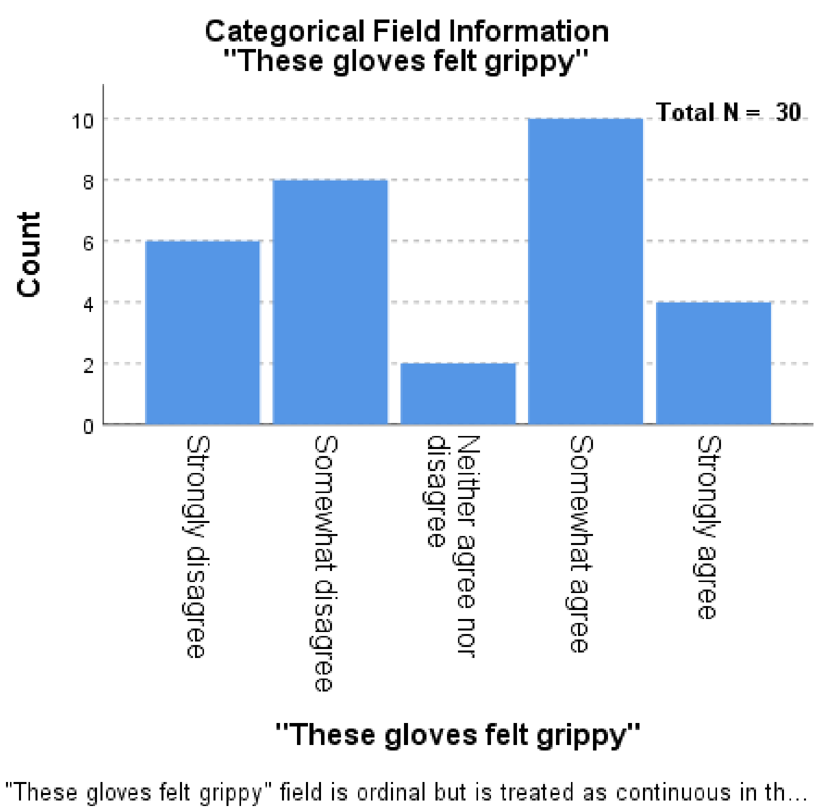
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
USG-USG + VT	-4.167	4.913	-.848	.396	1.000
USG-USG Nilo	-10.000	4.913	-2.035	.042	.627
USG-USG + 50% TPU	-15.000	4.913	-3.053	.002	.034
USG-USG + 100% TPU	-19.167	4.913	-3.901	.000	.001
USG + VT-USG Nilo	5.833	4.913	1.187	.235	1.000
USG + VT-USG + 50% TPU	10.833	4.913	2.205	.027	.412
USG + VT-USG + 100% TPU	15.000	4.913	3.053	.002	.034
USG Nilo-USG + 50% TPU	-5.000	4.913	-1.018	.309	1.000
USG Nilo-USG + 100% TPU	-9.167	4.913	-1.866	.062	.931
USG + 50% TPU-USG + 100% TPU	-4.167	4.913	-.848	.396	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.


Pairwise Comparisons of Glove
USG
15.83
USG + 100% TPU
10.00
USG + 50% TPU
25.00
No glove
5.83
USG Nilo
20.83
Each node shows the sample average rank of Glove.


Categorical Field Information
"These gloves felt grippy"
Total N = 30
Count
10
8
6
4
2
0
Strongly disagree
Somewhat disagree
Neither agree nor disagree
Somewhat agree
Strongly agree
"These gloves felt grippy"
"These gloves felt grippy" field is ordinal but is treated as continuous in th...

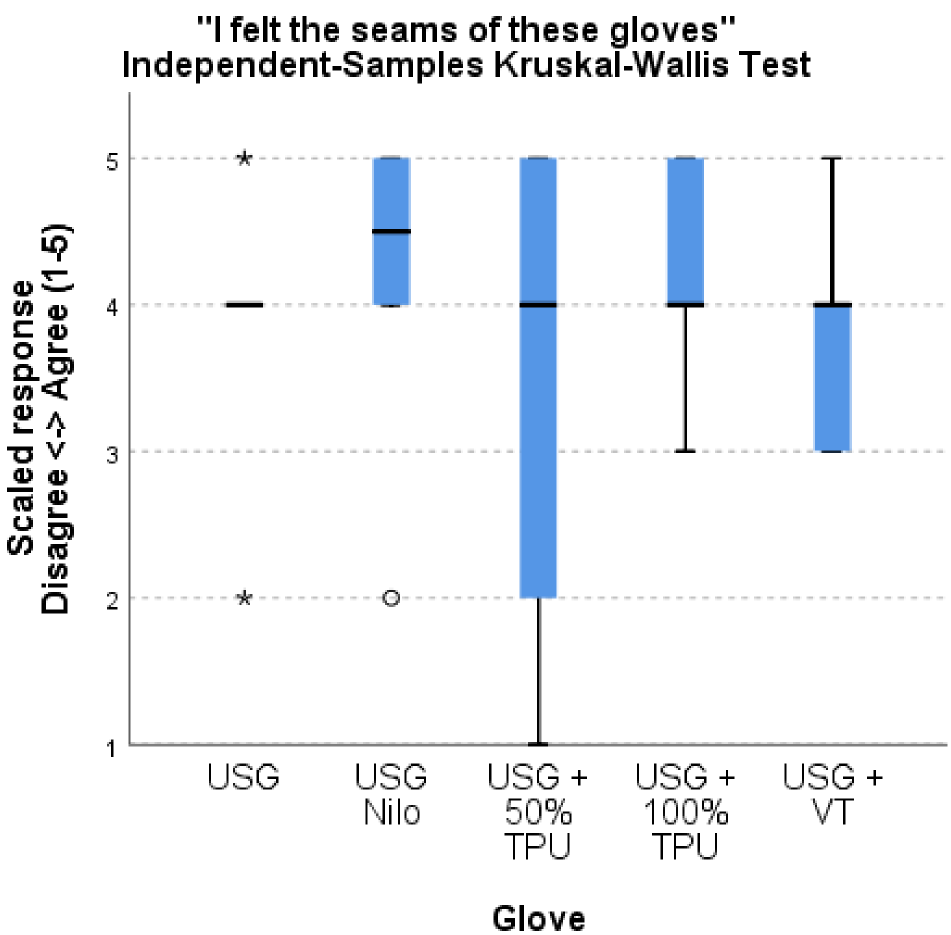
D.4.8 "I felt the seams of these gloves"

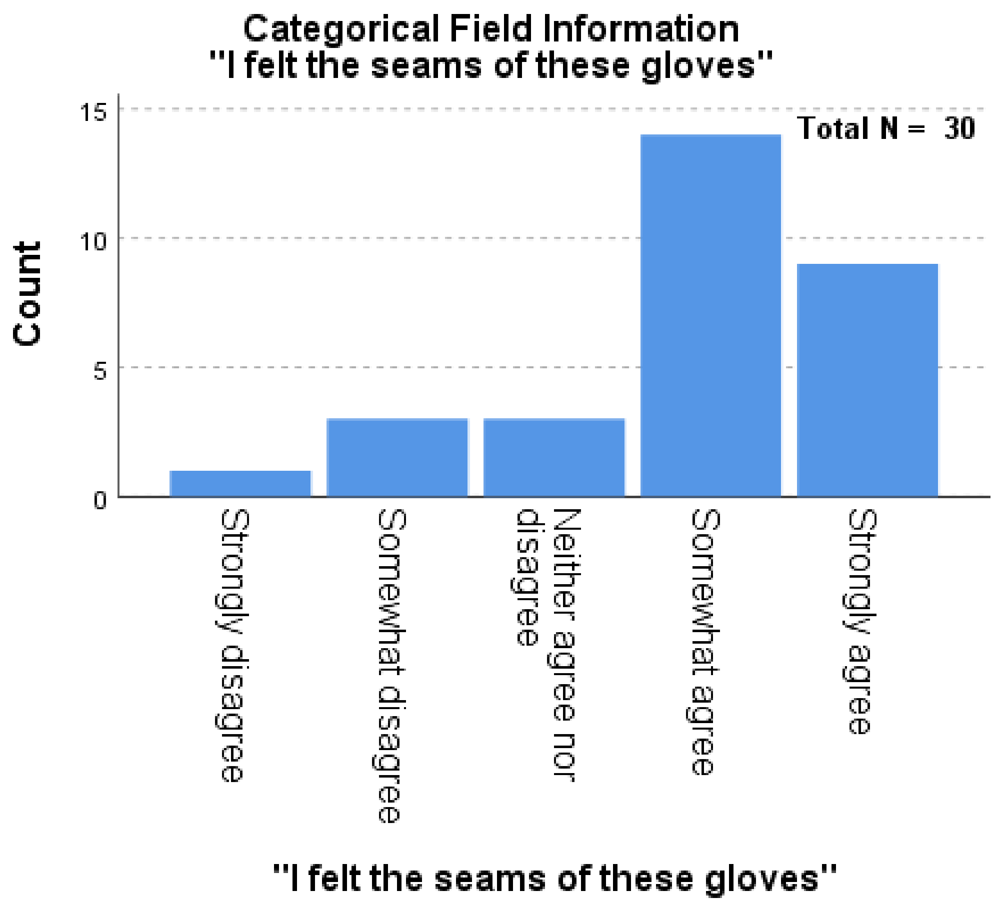
Independent-Samples Kruskal-Wallis Test Summary

Total N	30
Test Statistic	1.467[a,b]
Degree Of Freedom	4
Asymptotic Sig.(2-sided test)	.833

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.



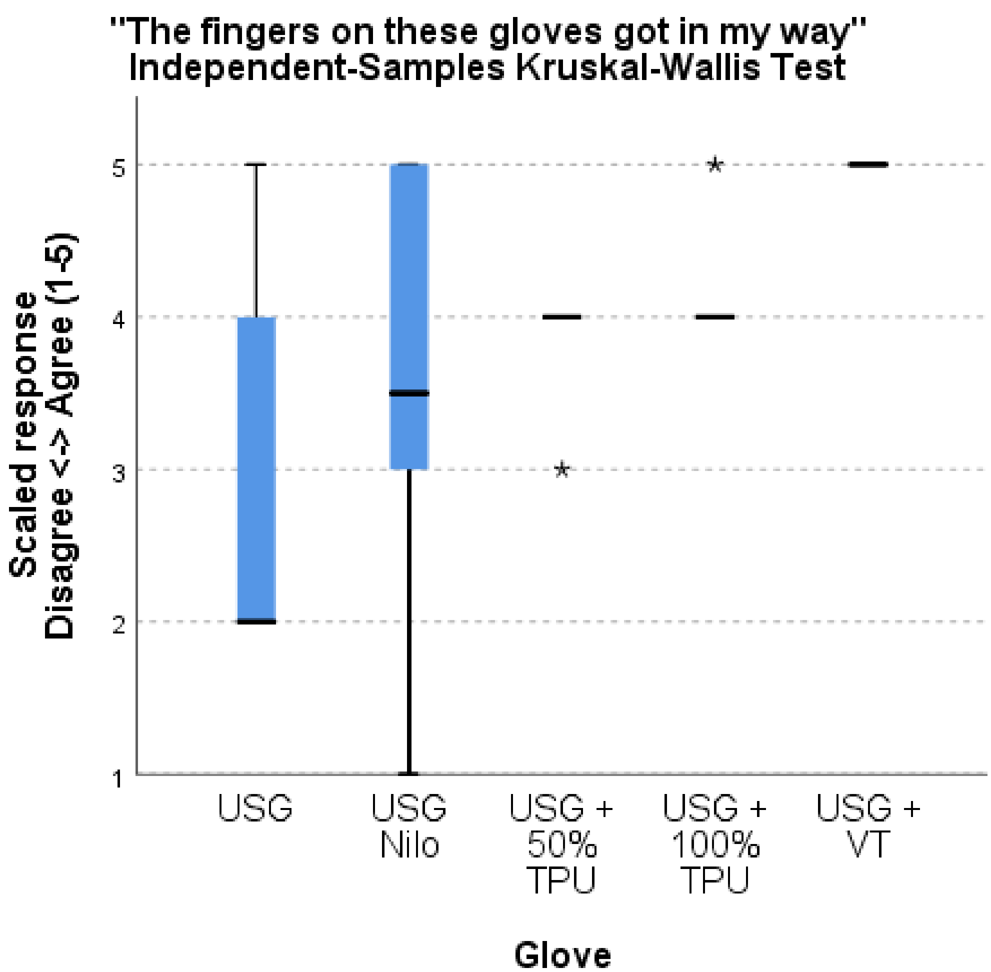


"I felt the seams of these gloves" field is ordinal but is treated as continuous in the test.

D.4.9 "The fingers on these gloves got in my way"

Independent-Samples Kruskal-Wallis Test Summary	
Total N	30
Test Statistic	13.156[a]
Degree Of Freedom	4
Asymptotic Sig.(2-sided test)	.011

a. The test statistic is adjusted for ties.



Pairwise Comparisons of Glove

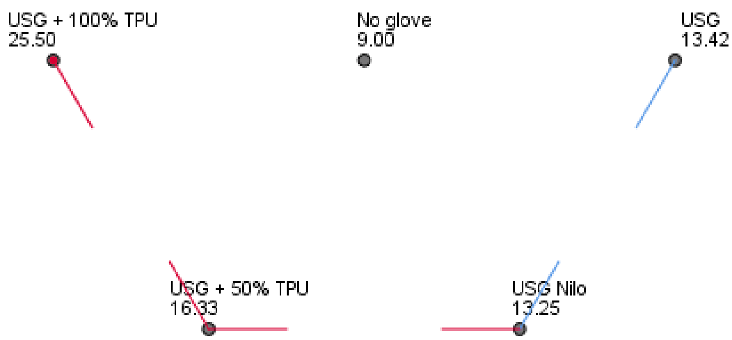
Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
USG-USG + 50% TPU	-4.250	4.813	-.883	.377	1.000
USG-USG Nilo	-4.417	4.813	-.918	.359	1.000
USG-USG + 100% TPU	-7.333	4.813	-1.524	.128	1.000
USG-USG + VT	-16.500	4.813	-3.429	.001	.009
USG + 50% TPU-USG Nilo	.167	4.813	.035	.972	1.000
USG + 50% TPU-USG + 100% TPU	-3.083	4.813	-.641	.522	1.000
USG + 50% TPU-USG + VT	-12.250	4.813	-2.545	.011	.164
USG Nilo-USG + 100% TPU	-2.917	4.813	-.606	.544	1.000
USG Nilo-USG + VT	-12.083	4.813	-2.511	.012	.181
USG + 100% TPU-USG + VT	-9.167	4.813	-1.905	.057	.852

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

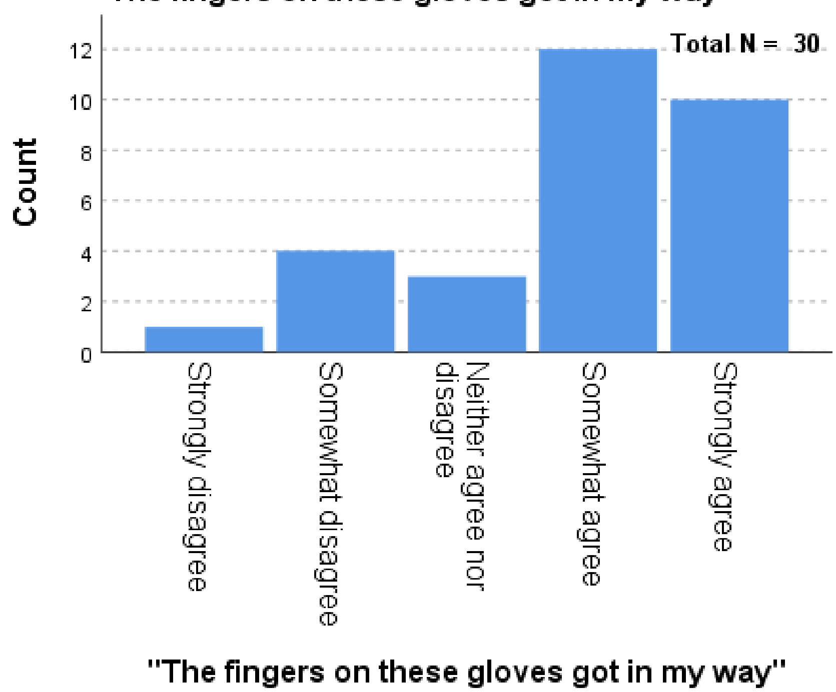
a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Pairwise Comparisons of Glove



Each node shows the sample average rank of Glove.

Categorical Field Information
"The fingers on these gloves got in my way"



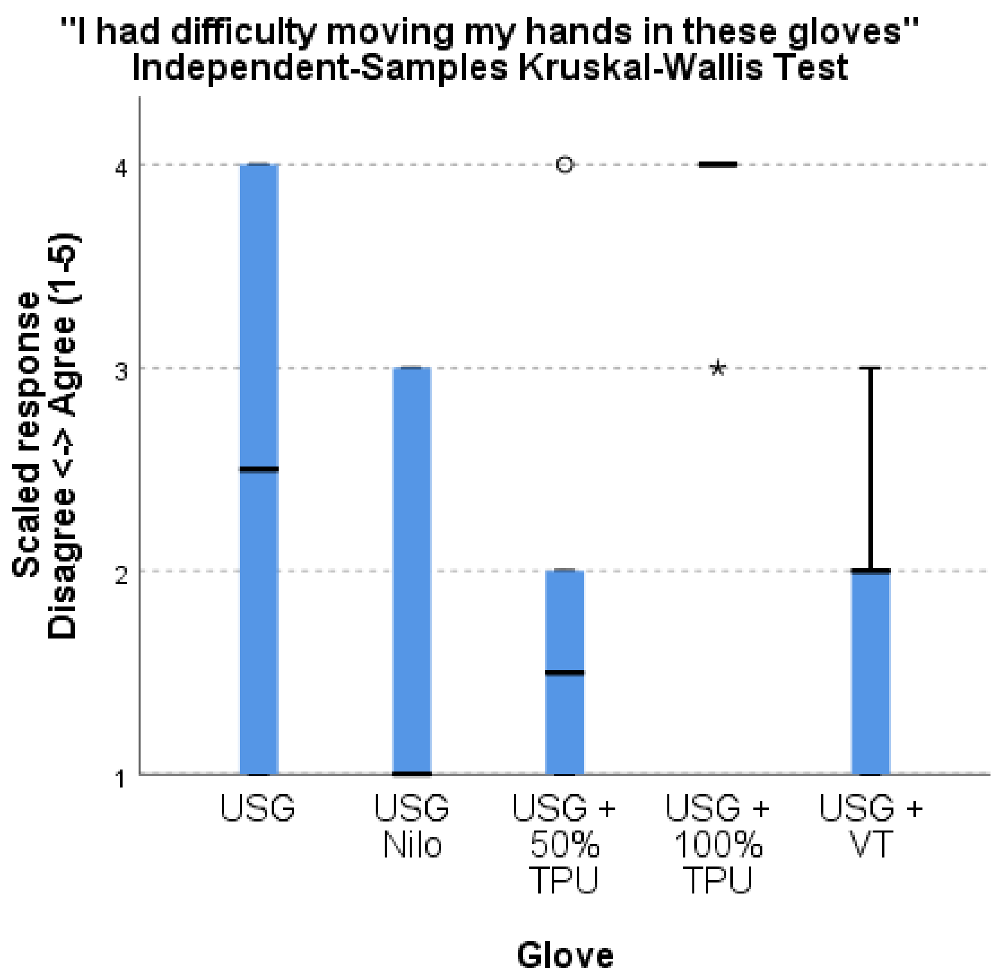
"The fingers on these gloves got in my way" field is ordinal but is treated as continuous in the test.

D.4.10 "I had difficulty moving my hands in these gloves" across Glove

Independent-Samples Kruskal-Wallis Test Summary

Total N	30
Test Statistic	11.943[a]
Degree Of Freedom	4
Asymptotic Sig.(2-sided test)	.018

a. The test statistic is adjusted for ties.



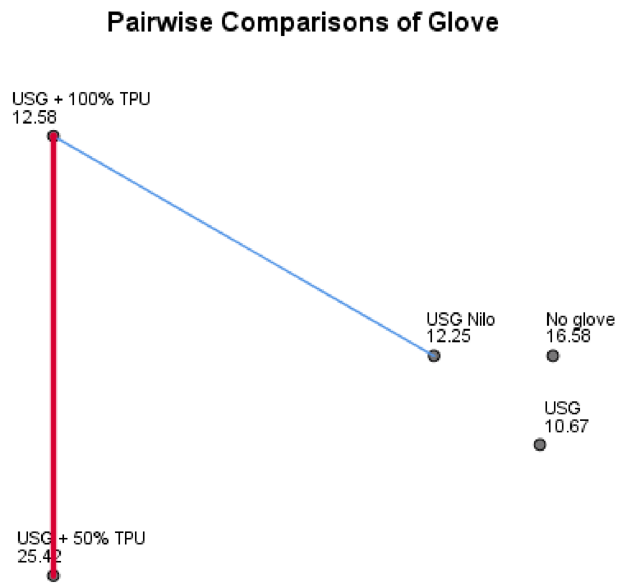
Pairwise Comparisons of Glove

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
USG Nilo-USG + 50% TPU	-1.583	4.875	-.325	.745	1.000
USG Nilo-USG + VT	-1.917	4.875	-.393	.694	1.000
USG Nilo-USG	5.917	4.875	1.214	.225	1.000
USG Nilo-USG + 100% TPU	-14.750	4.875	-3.025	.002	.037
USG + 50% TPU-USG + VT	-.333	4.875	-.068	.945	1.000
USG + 50% TPU-USG	4.333	4.875	.889	.374	1.000
USG + 50% TPU-USG + 100% TPU	-13.167	4.875	-2.701	.007	.104
USG + VT-USG	4.000	4.875	.820	.412	1.000
USG + VT-USG + 100% TPU	12.833	4.875	2.632	.008	.127
USG-USG + 100% TPU	-8.833	4.875	-1.812	.070	1.000

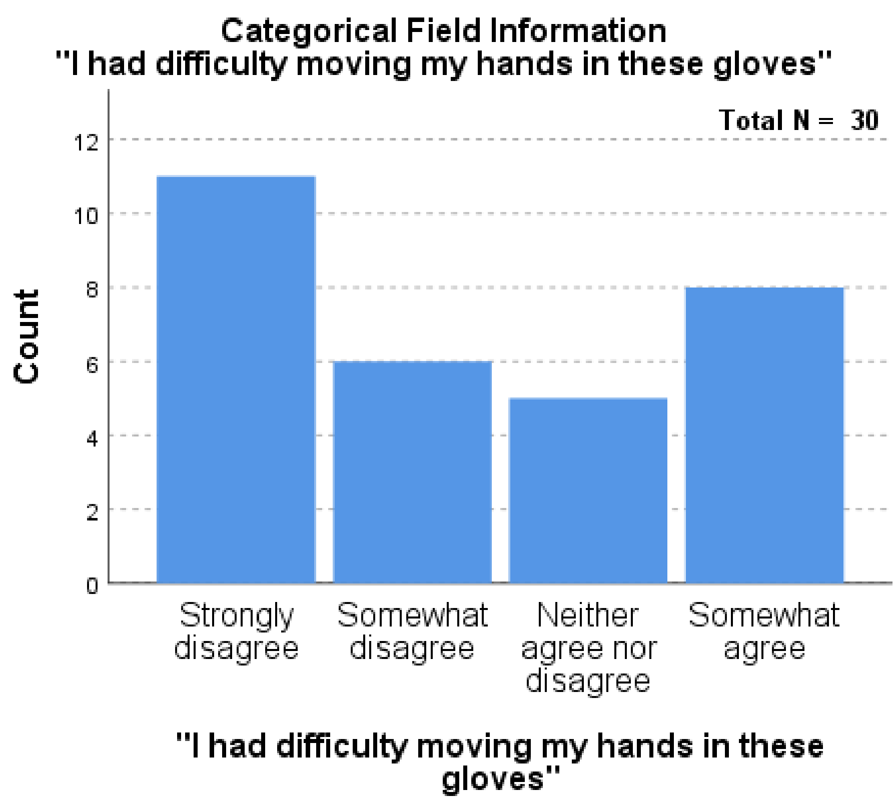
Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.

Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.



Each node shows the sample average rank of Glove.



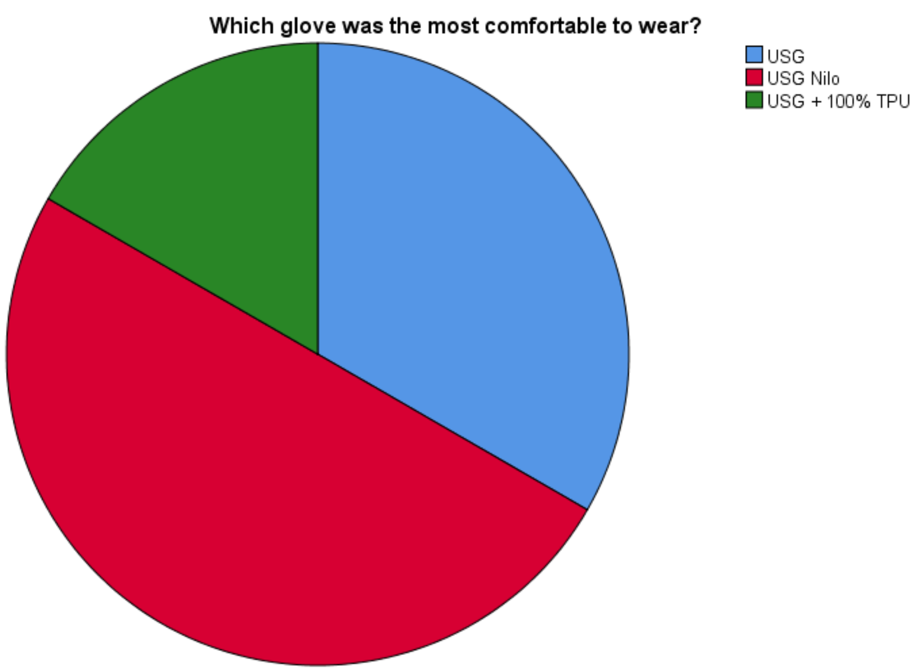
"I had difficulty moving my hands in these gloves" field is ordinal but is treated as continuous in the test.

D.5 Glove preference data

D.5.1 Which glove was the most comfortable to wear?

Which glove was the most comfortable to wear?

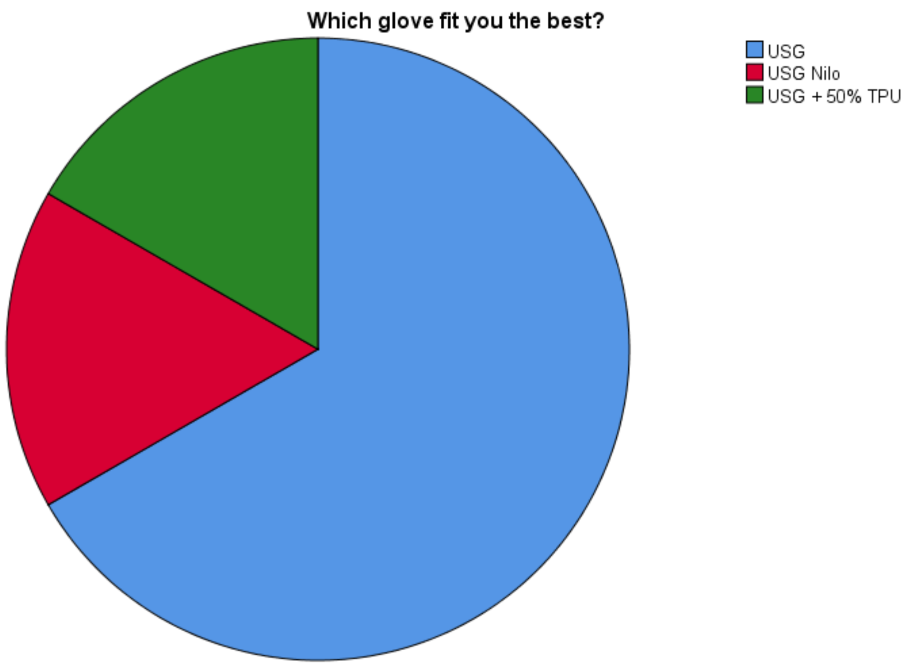
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	USG	2	33.3	33.3	33.3
	USG Nilo	3	50.0	50.0	83.3
	USG + 100% TPU	1	16.7	16.7	100.0
	Total	6	100.0	100.0	



D.5.2 Which glove fit you the best?

Which glove fit you the best?

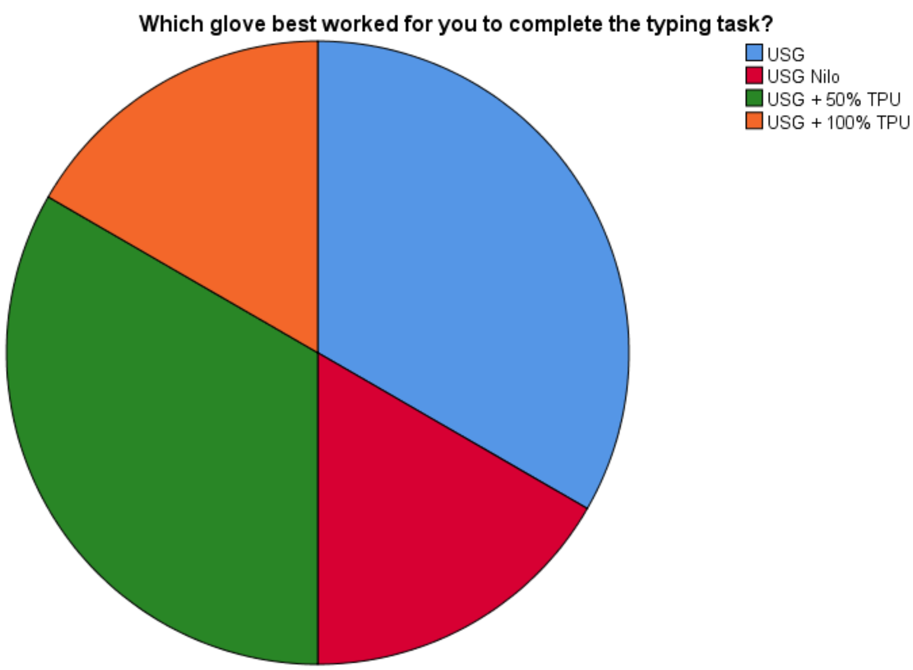
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	USG	4	66.7	66.7	66.7
	USG Nilo	1	16.7	16.7	83.3
	USG + 50% TPU	1	16.7	16.7	100.0
	Total	6	100.0	100.0	



D.5.3 Which glove best worked for you to complete the typing task?

Which glove best worked for you to complete the typing task?

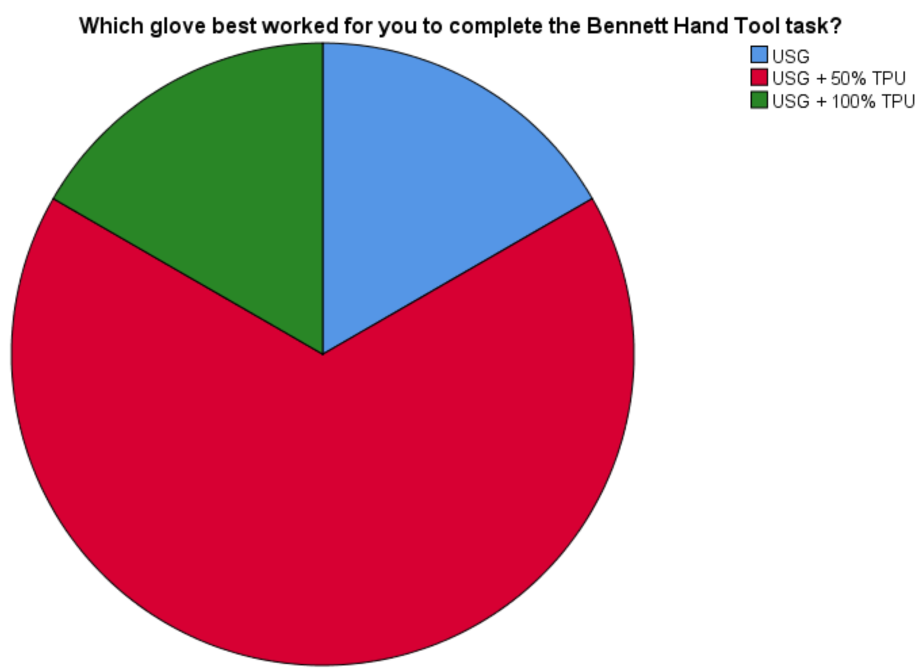
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	USG	2	33.3	33.3	33.3
	USG Nilo	1	16.7	16.7	50.0
	USG + 50% TPU	2	33.3	33.3	83.3
	USG + 100% TPU	1	16.7	16.7	100.0
	Total	6	100.0	100.0	



D.5.4 Which glove best worked for you to complete the Bennett Hand Tool task?

Which glove best worked for you to complete the Bennett Hand Tool task?

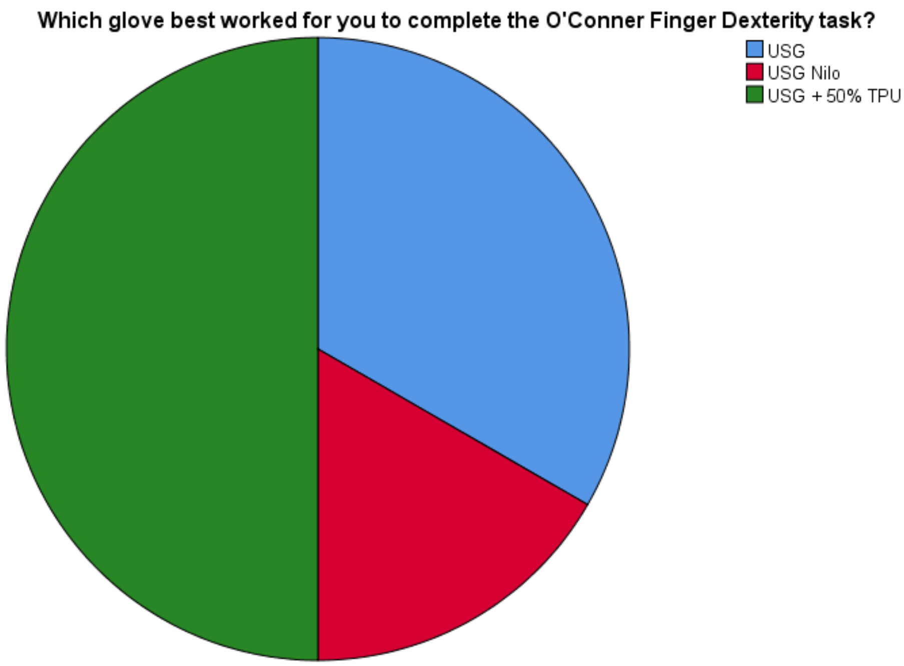
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	USG	1	16.7	16.7	16.7
	USG + 50% TPU	4	66.7	66.7	83.3
	USG + 100% TPU	1	16.7	16.7	100.0
	Total	6	100.0	100.0	



D.5.5 Which glove best worked for you to complete the O'Conner Finger Dexterity task?

Which glove best worked for you to complete the O'Conner Finger Dexterity task?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	USG	2	33.3	33.3	33.3
	USG Nilo	1	16.7	16.7	50.0
	USG + 50% TPU	3	50.0	50.0	100.0
	Total	6	100.0	100.0	



D.5.6 Which glove best worked for you to complete the Minnesota Manual Dexterity task?

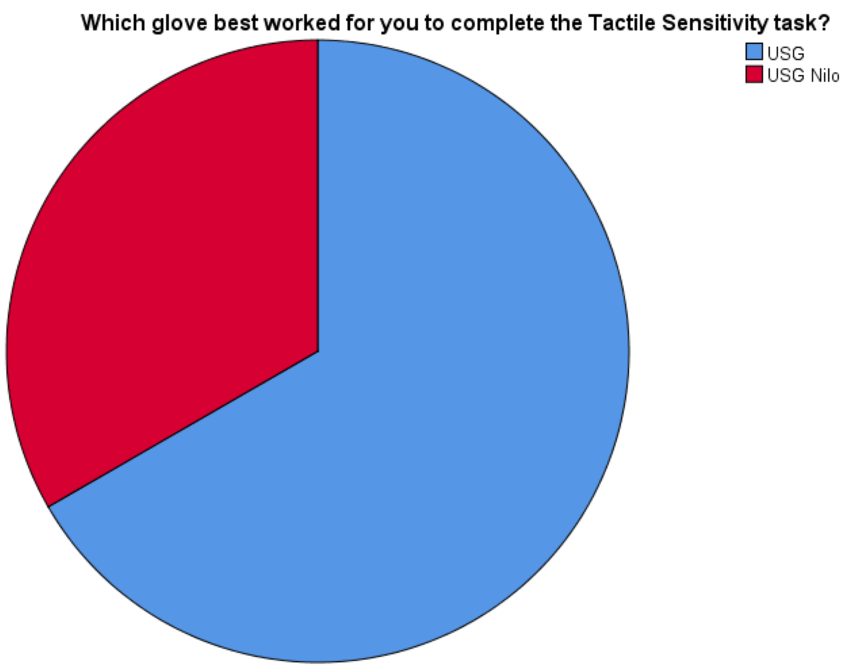
Which glove best worked for you to complete the Minnesota Manual Dexterity task?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	USG Nilo	1	16.7	16.7	16.7
	USG + 50% TPU	2	33.3	33.3	50.0
	USG + 100% TPU	3	50.0	50.0	100.0
	Total	6	100.0	100.0	

D.5.7 Which glove best worked for you to complete the Tactile Sensitivity task?

Which glove best worked for you to complete the Tactile Sensitivity task?

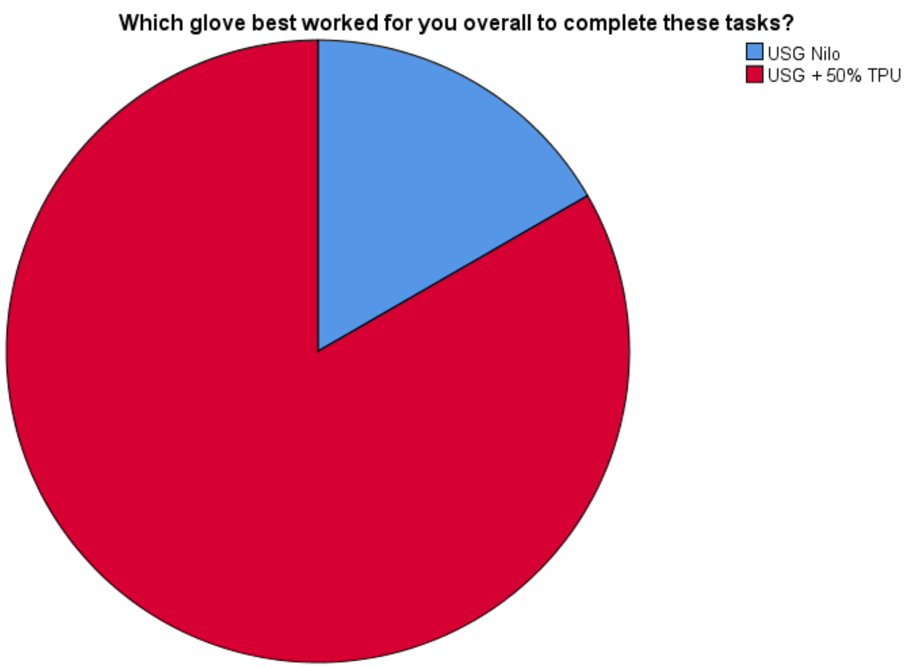
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	USG	4	66.7	66.7	66.7
	USG Nilo	2	33.3	33.3	100.0
	Total	6	100.0	100.0	



D.5.8 Which glove best worked for you overall to complete these tasks?

Which glove best worked for you overall to complete these tasks?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	USG Nilo	1	16.7	16.7	16.7
	USG + 50% TPU	5	83.3	83.3	100.0
	Total	6	100.0	100.0	



REFERENCES

1. Dorit H. Aaron and Caroline W. Stegink Jansen. 2003. Development of the Functional Dexterity Test (FDT): Construction, validity, reliability, and normative data. *Journal of Hand Therapy* 16, 1: 12–21. https://doi.org/10.1016/S0894-1130(03)80019-4

2. Farhang Akbar-Khanzadeh, Michael S. Bisesi, and Ruben D. Rivas. 1995. Comfort of personal protective equipment. *Applied Ergonomics* 26, 3: 195–198. https://doi.org/10.1016/0003-6870(95)00017-7

3. E. L. Altschuler, S. B. Wisdom, L. Stone, C. Foster, D. Galasko, D. M. Llewellyn, and V. S. Ramachandran. 1999. Rehabilitation of hemiparesis after stroke with a mirror. *The Lancet* 353: 2035–2036.

4. Godfried Augenbroe. 2011. The role of simulation in performance based building. In *Building performance simulation for design and operation*. 15–36. Retrieved from http://bit.ly/building-performance-simulation

5. Sandra F Bassett and Keith J Petrie. 1999. The Effect of Treatment Goals on Patient Compliance with Physiotherapy Exercise Programmes. *Physiotherapy* 85, 3: 130–137. https://doi.org/10.1016/S0031-9406(05)65693-3

6. Judith A. Bell-Krotoski, Elaine Ewing Fess, John H. Figarola, and Danell Hiltz. 1995. Threshold Detection and Semmes-Weinstein Monofilaments. *Journal of Hand Therapy* 8, 2: 155–162. https://doi.org/10.1016/S0894-1130(12)80314-0

7. Judith Bell-Krotoski, Sidney Weinstein, and Curt Weinstein. 1993. Testing Sensibility, Including Touch-Pressure, Two-point Discrimination, Point Localization, and Vibration. *Journal of Hand Therapy* 6, 2: 114–123. https://doi.org/10.1016/S0894-1130(12)80292-4

8. George K Bennett. 1965. *Hand-tool dexterity test*. Psychological Corporation.

9. Monique A. M. Berger, Arno J. Krul, and Hein A. M. Daanen. 2009. Task specificity of finger dexterity tests. *Applied Ergonomics* 40, 1: 145–147. https://doi.org/10.1016/j.apergo.2008.01.014

10. Ram R. Bishu, Sanjay Batra, David J. Cochran, and Michael W. Riley. 1987. Glove Effect on Strength: An Investigation of Glove Attributes. *Proceedings of the Human Factors Society Annual Meeting* 31, 8: 901–905. https://doi.org/10.1177/154193128703100812

11. Erin Blakemore. 2018. What Gloves Meant to the Victorians. *JSTOR Daily*. Retrieved August 17, 2019 from https://daily.jstor.org/what-gloves-meant-to-the-victorians/

12. Amy Boos, Qinyin Qiu, Gerard G. Fluet, and Sergei V. Adamovich. 2011. Haptically facilitated bimanual training combined with augmented visual feedback in moderate to severe hemiplegia. 3111–3114. Retrieved from ://WOS:000298810002209

13. N. N. Byl and G. M. Abrams. 2002. EFFECTIVENESS OF NEUROREHABILITATION BASED ON THE PRINCIPLES OF NEUROPLASTICITY:

PATIENTS SIX MONTHS POST STROKE. *Society for Neuroscience Abstract Viewer and Itinerary Planner* 2002: Abstract No. 650.5-Abstract No. 650.5.

14. A. Cacchio, E. De Blasis, S. Necozione, F. di Orio, and V. Santilli. 2009. Mirror Therapy for Chronic Complex Regional Pain Syndrome Type 1 and Stroke. *New England Journal of Medicine* 361, 6: 634–636. https://doi.org/10.1056/NEJMc0902799

15. Milène Catoire, Bouke N. Krom, and Jan B. F. van Erp. 2018. Towards a Test Battery to Benchmark Dexterous Performance in Teleoperated Systems. In *Haptics: Science, Technology, and Applications* (Lecture Notes in Computer Science), 440–451. https://doi.org/10.1007/978-3-319-93399-3_38

16. British Design Council. 2005. *The "double diamond" design process model*. Retrieved from http://www.designcouncil.org.uk/designprocess

17. John W. Creswell and Vicki L. Plano Clark. 2018. *Designing and conducting mixed methods research*. SAGE, Los Angeles.

18. Johanne Desrosiers, Annie Rochette, Réjean Hébert, and Gina Bravo. 1997. The Minnesota Manual Dexterity Test: Reliability, Validity and Reference Values Studies with Healthy Elderly People. *Canadian Journal of Occupational Therapy* 64, 5: 270–276. https://doi.org/10.1177/000841749706400504

19. Iman Dianat, Christine M. Haslegrave, and Alex W. Stedmon. 2012. Methodology for evaluating gloves in relation to the effects on hand performance capabilities: a literature review. *Ergonomics* 55, 11: 1429–1451. https://doi.org/10.1080/00140139.2012.708058

20. Iman Dianat, Christine M. Haslegrave, and Alex W. Stedmon. 2014. Design options for improving protective gloves for industrial assembly work. *Applied Ergonomics* 45, 4: 1208–1217. https://doi.org/10.1016/j.apergo.2014.02.009

21. Norman Doidge. 2007. *The brain that changes itself: Stories of personal triumph from the frontiers of brain science*. Penguin.

22. P. Duncan, L. Richards, D. Wallace, J. Stoker-Yates, P. Pohl, C. Luchies, A. Ogle, and S. Studenski. 1998. A randomized, controlled pilot study of a home-based exercise program for individuals with mild and moderate stroke. *Stroke* 29, 10: 2055–2060. https://doi.org/10.1161/01.str.29.10.2055

23. Elaine Fess. 2002. Documentation: Essential elements of an upper extremity assessment battery. In *Rehabilitation of the Hand*, A. L. Osterman, E. Mackin, A. Callahan and T. Skirvin (eds.). 263–284.

24. Francine Gemperle, Chris Kasabach, John Stivoric, Malcolm Bauer, and Richard Martin. 1998. Design for wearability. In *Wearable Computers, 1998. Digest of Papers. Second International Symposium on*, 116–122. https://doi.org/10.1109/ISWC.1998.729537

25. Guido Gioberto and Lucy Dunne. 2012. Theory and characterization of a top-thread coverstitched stretch sensor. In *2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 3275–3280. https://doi.org/10.1109/ICSMC.2012.6378296

26. Guido Gioberto and Lucy E. Dunne. 2012. Garment Positioning and Drift in Garment-Integrated Wearable Sensing. In *2012 16th International Symposium on Wearable Computers*, 64–71. https://doi.org/10.1109/ISWC.2012.35

27. V. Gnaneswaran, B. Mudhunuri, and R. R. Bishu. 2008. A study of latex and vinyl gloves: Performance versus allergy protection properties. *International Journal of Industrial Ergonomics* 38, 2: 171–181. https://doi.org/10.1016/j.ergon.2007.10.027

28. C. James Goodwin and Kerri A. Goodwin. 2013. *Research in psychology: methods and design*. Wiley, Hoboken, NJ.

29. O. Grobnicu, F. Vinée, Y. Igeta, F. Xavier, T. Bourcier, and P. Liverneaux. 2018. Tactile sensitivity thresholds for the radial hemi-pulp of the index: A comparison between the Semmes-Weinstein and Cochet-Bonnet tests in 25 healthy subjects. *Hand Surgery and Rehabilitation* 37, 5: 295–299. https://doi.org/10.1016/j.hansur.2018.07.003

30. Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, Peter A. Hancock and Najmedin Meshkati (eds.). North-Holland, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

31. Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology*, Peter A. Hancock and Najmedin Meshkati (eds.). North-Holland, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

32. Kate Hartman, Brian Jepson, Emma Dvorak, and Rebecca Demarest. 2014. *Wearable electronics*. Maker Media, Inc., Sebastopol, CA.

33. George Havenith, Ingvar Holmér, and Ken Parsons. 2002. Personal factors in thermal comfort assessment: clothing properties and metabolic heat production. *Energy and Buildings* 34, 6: 581–591. https://doi.org/10.1016/S0378-7788(02)00008-7

34. I. Herbst and J. Stark. 2005. Comparing force magnitudes by means of vibro-tactile, auditory, and visual feedback. In *IEEE International Workshop on Haptic Audio Visual Environments and their Applications*, 5 pp.-. https://doi.org/10.1109/HAVE.2005.1545654

35. Franck Hernoux and Olivier Christmann. 2015. A seamless solution for 3D real-time interaction: design and evaluation. *Virtual Reality* 19, 1: 1–20. https://doi.org/10.1007/s10055-014-0255-z

36. Ching-Tang Huang, Chien-Lung Shen, Chien-Fa Tang, and Shuo-Hung Chang. 2008. A wearable yarn-based piezo-resistive sensor. *Sensors and Actuators A: Physical* 141, 2: 396–403. https://doi.org/10.1016/j.sna.2007.10.069

37. IEEE/RSJ International Conference on Intelligent Robots and Systems. 2017. Robotic Grasping and Manipulation Competition @ IROS 2017. *IEEE RAS Technical Committee Robotic Hand Grasping and Manipulation*. Retrieved November 19, 2020 from http://www.rhgm.org/activities/competition_iros2017/

38. ISO 9283:1998. 1998. *Manipulating industrial robots — Performance criteria and related test methods*. ISO/TC 299 Robotics. Retrieved November 19, 2020 from

https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/02/22/22244.html

39. D. Jack, R. Boian, A. S. Merians, M. Tremaine, G. C. Burdea, S. V. Adamovich, M. Recce, and H. Poizner. 2001. Virtual reality-enhanced stroke rehabilitation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 9, 3: 308–318. https://doi.org/10.1109/7333.948460

40. Li Jiang, Mark R. Cutkosky, Juhani Ruutiainen, and Roope Raisamo. 2008. Improving finger force control with vibrational haptic feedback for Multiple Sclerosis. 110–115.

41. B. B. Johansson. 2012. Multisensory stimulation in stroke rehabilitation. *Frontiers in Human Neuroscience* 6. https://doi.org/10.3389/fnhum.2012.00060

42. Myounggu Kang. 2004. 11.220 Quantitative Reasoning and Statistical Methods for Planning (MIT/DUSP). Retrieved October 13, 2020 from http://web.mit.edu/11.220/www05/brushup/spss/boxplot.htm

43. James F. Knight and Chris Baber. 2007. Assessing the physical loading of wearable computers. *Applied Ergonomics* 38, 2: 237–247. https://doi.org/10.1016/j.apergo.2005.12.008

44. M. M. Lee, H. Y. Cho, and C. H. Song. 2012. The Mirror Therapy Program Enhances Upper-Limb Motor Recovery and Motor Function in Acute Stroke Patients.

American Journal of Physical Medicine & Rehabilitation 91, 8: 689–696. https://doi.org/10.1097/PHM.0b013e31824fa86d

45. Jeff Lieberman and Cynthia Breazeal. 2007. TIKL: Development of a Wearable Vibrotactile Feedback Suit for Improved Human Motor Learning. *IEEE Transactions on Robotics* 23, 5: 919–926. https://doi.org/10.1109/TRO.2007.907481

46. Pär Linnertz, Johanna Prieto Ek, and Birgitta Rosén. 2018. Shape-texture-identification—STI—A test for tactile gnosis: Concurrent validity of STI2. *Journal of Hand Therapy*. https://doi.org/10.1016/j.jht.2018.05.004

47. LiveChat. Free Typing Test - Check Your Typing Speed in 60 Seconds | LiveChat Tools. *LiveChat*. Retrieved November 2, 2020 from https://www.livechat.com/typing-speed-test/

48. Dave Madison. 2017. Repurposing the Staples Easy Button. *Parts Not Included*. Retrieved September 20, 2020 from https://www.partsnotincluded.com/repurposing-the-staples-easy-button/

49. Eliane C. Magdalon, Stella M. Michaelsen, Antonio A. Quevedo, and Mindy F. Levin. 2011. Comparison of grasping movements made by healthy subjects in a 3-dimensional immersive virtual versus physical environment. *Acta Psychologica* 138, 1: 126–134. https://doi.org/10.1016/j.actpsy.2011.05.015

50. Hernando Gonzalez Malabet, Rafael Alvarez Robles, and Kyle B. Reed. 2010. Symmetric Motions for Bimanual Rehabilitation. In *Ieee/Rsj 2010 International*

Conference on Intelligent Robots and Systems. 5133–5138. Retrieved from ://WOS:000287672004108

51. Tanya Thais Markow. 2012. Mobile music touch: using haptic stimulation for passive rehabilitation and learning. Georgia Institute of Technology. Retrieved September 1, 2020 from https://smartech.gatech.edu/handle/1853/43665

52. V. Mathiowetz, G. Volland, N. Kashman, and K. Weber. 1985. Adult norms for the Box and Block Test of manual dexterity. *The American journal of occupational therapy. : official publication of the American Occupational Therapy Association* 39, 6: 386–391. https://doi.org/10.5014/ajot.39.6.386

53. K. Matthys, M. Smits, J. N. Van der Geest, A. Van der Lugt, R. Seurinck, H. J. Stam, and R. W. Selles. 2009. Mirror-Induced Visual Illusion of Hand Movements: A Functional Magnetic Resonance Imaging Study. *Archives of Physical Medicine and Rehabilitation* 90, 4: 675–681. https://doi.org/10.1016/j.apmr.2008.09.571

54. Samuel McAmis and Kyle B. Reed. 2011. Symmetry Modes and Stiffnesses for Bimanual Rehabilitation. In *2011 Ieee International Conference on Rehabilitation Robotics*. Retrieved from ://WOS:000299169800169

55. C. McCabe. 2011. Mirror Visual Feedback Therapy. A Practical Approach. *Journal of Hand Therapy* 24, 2: 170–178. https://doi.org/10.1016/j.jht.2010.08.003

56. Stuart McFarlane and Clement Fletcher. 2009. How to Make Anything (Using Acrylic and Machine Screws). *Instructables*. Retrieved September 20, 2020 from

https://www.instructables.com/id/How-to-Make-Anything-Using-Acrylic-and-Machine-Sc/

57. A. S. Merians, E. Tunik, G. G. Fluet, Q. Qiu, and S. V. Adamovich. 2009. Innovative approaches to the rehabilitation of upper extremity hemiparesis using virtual environments. *European Journal of Physical and Rehabilitation Medicine* 45, 1: 123–133.

58. Merriam-Webster.com. Encumbrance. *Merriam-Webster.com Dictionary*. Retrieved November 2, 2020 from https://www.merriam-webster.com/dictionary/encumbrance

59. Anil Mital, Tsuang Kuo, and Hamid F Faard. 1994. A quantitative evaluation of gloves used with non-powered hand tools in routine maintenance tasks. *Ergonomics* 37, 2: 333–343. https://doi.org/10.1080/00140139408963650

60. A. Muralidhar, R. R. Bishu, and M. S. Hallbeck. 1999. The development and evaluation of an ergonomic glove. *Applied Ergonomics* 30, 6: 555–563. https://doi.org/10.1016/S0003-6870(99)00005-8

61. Damien Newman. *Design Process Squiggle*. Retrieved from http://v2.centralstory.com/about/squiggle/

62. I. Nojima, T. Mima, S. Koganemaru, M. N. Thabit, H. Fukuyama, and T. Kawamata. 2012. Human Motor Plasticity Induced by Mirror Visual Feedback. *Journal of Neuroscience* 32, 4: 1293–1300. https://doi.org/10.1523/jneurosci.5364-11.2012

63. D. Odell and E. Faggin. 2014. The typing performance and preference costs of reducing tactile feedback and tactile landmarks in tablet keyboards. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 58, 1: 1790–1794. https://doi.org/10.1177/1541931214581373

64. Margaret Orth. 2002. Defining Flexibility and Sewability in Conductive Yarns. *MRS Online Proceedings Library Archive* 736. https://doi.org/10.1557/PROC-736-D1.4

65. Fred G. W. C. Paas, Jeroen J. G. van Merriënboer, and Jos J. Adam. 1994. Measurement of Cognitive Load in Instructional Research. *Perceptual and Motor Skills* 79, 1: 419–430. https://doi.org/10.2466/pms.1994.79.1.419

66. Debajyoti Pati. 2011. A Framework for Evaluating Evidence in Evidence-Based Design. *HERD: Health Environments Research & Design Journal* 4, 3: 50–71. https://doi.org/10.1177/193758671100400305

67. Hannah Perner-Wilson, Leah Buechley, and Mika Satomi. 2010. Handcrafting textile interfaces from a kit-of-no-parts. In *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction* (TEI ’11), 61–68. https://doi.org/10.1145/1935701.1935715

68. E. R. Post, M. Orth, P. R. Russo, and N. Gershenfeld. 2000. E-broidery: Design and fabrication of textile-based computing. *IBM Systems Journal* 39, 3.4: 840–860. https://doi.org/10.1147/sj.393.0840

69. Michael Power. 2004. After a stroke, ability with daily tasks of living improves after therapy based rehabilitation services. *Evidence-based Healthcare* 8, 4: 188–189. https://doi.org/10.1016/j.ehbc.2004.05.018

70. Halley P. Profita, James Clawson, Scott Gilliland, Clint Zeagler, Thad Starner, Jim Budd, and Ellen Yi-Luen Do. 2013. Don'T Mind Me Touching My Wrist: A Case Study of Interacting with On-body Technology in Public. In *Proceedings of the 2013 International Symposium on Wearable Computers* (ISWC '13), 89–96. https://doi.org/10.1145/2493988.2494331

71. Rahul Rajan, Ted Selker, and Ian Lane. 2016. Task Load Estimation and Mediation Using Psycho-physiological Measures. In *Proceedings of the 21st International Conference on Intelligent User Interfaces* (IUI '16), 48–59. https://doi.org/10.1145/2856767.2856769

72. V. S. Ramachandran and E. L. Altschuler. 2009. The use of visual feedback, in particular mirror visual feedback, in restoring brain function. *Brain* 132: 1693–1710. https://doi.org/10.1093/brain/awp135

73. P. Richard and P. Coiffet. 1999. Dextrous haptic interaction in virtual environments: human performance evaluations. In *8th IEEE International Workshop on Robot and Human Interaction. RO-MAN '99 (Cat. No.99TH8483)*, 315–320. https://doi.org/10.1109/ROMAN.1999.900359

74. Joshua S. Rubinstein, David E. Meyer, and Jeffrey E. Evans. 2001. Executive control of cognitive processes in task switching. *Journal of experimental psychology: human perception and performance* 27, 4: 763.

75. Jeff Sauro and Joseph S. Dumas. 2009. Comparison of Three One-question, Post-task Usability Questionnaires. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '09), 1599–1608. https://doi.org/10.1145/1518701.1518946

76. Jeff Sauro and James R. Lewis. 2016. *Quantifying the user experience: practical statistics for user research*. Morgan Kaufmann, Cambridge.

77. Eric H. Schumacher, Travis L. Seymour, Jennifer M. Glass, David E. Fencsik, Erick J. Lauber, David E. Kieras, and David E. Meyer. 2001. Virtually Perfect Time Sharing in Dual-Task Performance: Uncorking the Central Cognitive Bottleneck. *Psychological Science* 12, 2: 101–108. https://doi.org/10.1111/1467-9280.00318

78. Caitlyn Seim, James Hallam, Shashank Raghu, Tri-An Le, Greg Bishop, and Thad Starner. 2015. *Perception in Hand-Worn Haptics: Placement, Simultaneous Stimuli, and Vibration Motor Comparisons*. Georgia Institute of Technology. Retrieved November 2, 2020 from https://smartech.gatech.edu/handle/1853/55952

79. Staples, Inc. Staples® Easy Button™ at Staples. *Staples.com*. Retrieved September 20, 2020 from /Staples-Easy-Button/product_606396

80. S. Sutbeyaz, G. Yavuzer, N. Sezer, and B. F. Koseoglu. 2007. Mirror therapy enhances lower-extremity motor recovery and motor functioning after stroke: A randomized controlled trial. *Archives of Physical Medicine and Rehabilitation* 88, 5: 555–559. https://doi.org/10.1016/j.apmr.2007.02.034

81. J. Takasugi, D. Matsuzawa, T. Murayama, K. Nakazawa, K. Numata, and E. Shimizu. 2011. Referred sensations induced by a mirror box in healthy subjects. *Psychological Research-Psychologische Forschung* 75, 1: 54–60. https://doi.org/10.1007/s00426-010-0287-2

82. R. M. TAYLOR and J. V. F. BERMAN. 1982. Ergonomic aspects of aircraft keyboard design: the effects of gloves and sensory feedback on keying performance. *Ergonomics* 25, 11: 1109–1123. https://doi.org/10.1080/00140138208925068

83. Joseph Tiffin and E. J. Asher. 1948. The Purdue Pegboard: norms and studies of reliability and validity. *Journal of Applied Psychology* 32, 3: 234–247. https://doi.org/10.1037/h0061266

84. George E Torrens and Anne Newman. The evaluation of gloved and ungloved hands. *Contemporary ergonomics*: 301–305.

85. Elisabeth H. Tracey, Aaron J. Greene, and Richard L. Doty. 2012. Optimizing reliability and sensitivity of Semmes–Weinstein monofilaments for establishing point tactile thresholds. *Physiology & Behavior* 105, 4: 982–986. https://doi.org/10.1016/j.physbeh.2011.11.002

86. Manuel Veit, Antonio Capobianco, and Dominique Bechmann. 2008. Consequence of Two-handed Manipulation on Speed, Precision and Perception on Spatial Input Task in 3D Modelling Applications. *Journal of Universal Computer Science* 14, 19: 3174–3187.

87. Susan M. Watkins and Lucy E. Dunne. 2015. *Functional clothing design: from sportswear to spacesuits*. Fairchild Books, an imprint of Bloomsbury Publishing Inc, New York.

88. R. Wells, S. Hunt, K. Hurley, and P. Rosati. 2010. Laboratory assessment of the effect of heavy rubber glove thickness and sizing on effort, performance and comfort. *International Journal of Industrial Ergonomics* 40, 4: 386–391. https://doi.org/10.1016/j.ergon.2010.03.002

89. Kirsten Willms, Richard Wells, and Heather Carnahan. 2009. Glove Attributes and Their Contribution to Force Decrement and Increased Effort in Power Grip. *Human Factors* 51, 6: 797–812. https://doi.org/10.1177/0018720809357560

90. Gunes Yavuzer, Ruud Selles, Nebahat Sezer, Serap Sütbeyaz, Johannes B. Bussmann, Füsun Köseoğlu, Mesut B. Atay, and Henk J. Stam. 3. Mirror Therapy Improves Hand Function in Subacute Stroke: A Randomized Controlled Trial. *Archives of Physical Medicine and Rehabilitation* 89, 3: 393–398. https://doi.org/10.1016/j.apmr.2007.08.162

91. Clint Zeagler. 2017. Where to Wear It: Functional, Technical, and Social Considerations in On-body Location for Wearable Technology 20 Years of Designing for

Wearability. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers* (ISWC ’17), 150–157. https://doi.org/10.1145/3123021.3123042